DOCTORAL THESIS

Thinking Like an Expert Lawyer: Measuring Specialist Legal Expertise Through Think-Aloud Problem Solving and Verbal Protocol Analysis

MacMillan, Peter

Award date:
2015

THINKING LIKE AN EXPERT LAWYER: MEASURING SPECIALIST LEGAL EXPERTISE THROUGH THINK-ALOUD PROBLEM SOLVING AND VERBAL PROTOCOL ANALYSIS

PETER JOHN MACMILLAN

A thesis submitted in total fulfilment of the requirements
of the degree of Doctor of Philosophy

FACULTY OF LAW, BOND UNIVERSITY

MAY 2015

THIS PAGE IS INTENTIONALLY BLANK

ABSTRACT

This empirical study identifies and measures ways in which legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts.

The 20 participants in this study are all specialists in competition law. They are a mixture of lawyers and economists from eight leading Australian law and economics firms, and from two regulatory bodies responsible for administering national competition laws. This sample of individuals, whose years of experience as competition law specialists range from five to 35 years, is assumed to include apprentices, journeymen, experts and masters according to the proficiency scale used in Middle Ages craft guilds and still used today in studies of expertise and expert performance.

Following an initial selection and ranking process, participants are categorised according to these four levels of proficiency. Their cognitive skills are then tested and compared to identify expertise-related differences. The principal data analysed are the think-aloud, concurrent verbalisations of study participants recorded as they seek to assess legal risk in test cases that require the use of competition law expertise. These data reveal a number of readily identifiable and measurable differences between how more and less expert participants assess legal risk in their common area of legal specialisation.

Master-level legal specialists are identified by their correct substantive analysis of legal issues, the ease and speed with which they identify key issues, and their heavy and effective reliance on intuition in assessing legal risk. Experts are identified by their refusals to provide concluded views when information and time are limited, their ability to identify key issues quickly, and their effective integration of intuition and analytical reasoning.

In terms of lower-level legal specialists, journeymen are identified by their reliance on superficial reasoning to assess legal risk, their laboured identification of issues, and their tendency to rely on guessing rather than intuition and to 'blurt out' their responses prematurely. Apprentices are identified by their laboured reasoning and extensive

searching for relevant issues and analogies, their heavy reliance on analytical reasoning, and the influence of self-doubt on their deliberations.

These and related findings, which are generalizable across other areas of legal specialisation, confirm the results of previous studies, raise questions about others, and offer new insights into the ways in which lawyers at different points along the legal-expertise continuum think differently from each other. These insights have the potential to improve the assessment methodologies used in lawyer accreditation schemes, change how users of legal services assess the expertise of legal specialists, increase the effectiveness of in-house training programs developed within law firms and by other legal service providers, and lead to new pedagogical approaches to the design and delivery of post-graduate, mid-career courses offered by law schools. They also provide a new foundation for scholarly research into the cognitive development of specialist legal expertise, which to date has not distinguished between the four levels of proficiency identified in this thesis.

DECLARATION

This thesis is submitted to Bond University in fulfilment of the requirements of the Degree of Doctor of Philosophy.

This thesis represents my own original work towards this research degree and contains no material which has been previously submitted for a degree or diploma at this University or any other institution, except where due acknowledgement is made.

Signature: …… Date: 6 May 2015

ACKNOWLEDGEMENTS

This thesis would not have been completed without the unshakeable support and encouragement of my wife, Stella, and daughter, Hayley.

Eugene Clark and Geof Stapledon graciously agreed to support my enrolment application. Their generosity of spirit after so many years still inspires me.

I could not have asked for a better supervisory team than Nick James and Peter Earl, both of whom provided the ideal balance of wise guidance and space to learn.

I am profoundly indebted to each of the 20 legal specialists who selflessly found time in their work schedules to contribute anonymously to a research concept.

The University of Queensland generously provided me with a scholarship to start this project, and after transferring my studies to maintain supervisor continuity, Bond University was equally generous in providing me with a stipend to complete it.

Peter Macmillan
May 2015

TABLE OF CONTENTS

TABLES

GRAPHS, CHARTS AND FIGURES

I INTRODUCTION

In its broadest setting, this thesis forms part of the scholarly literature devoted to understanding human wisdom and the nature of expertise. It investigates specific aspects of how lawyers think, which has for decades been a topic of keen interest amongst legal scholars yet also one that has been rarely subjected to empirical investigation. Previous studies in this area have overwhelmingly focused on the generic and knowledge-based aspects of legal thinking skills. They have also mainly been restricted to comparisons between novices and presumed legal experts. This study is the first to identify cognitive differences between legal specialists with different levels of expertise within the same area of law.

Using investigative methodologies from the field of cognitive psychology, this thesis is grounded in the traditional framework of progressive or relative expertise. This is a departure from previous studies that have relied on knowledge-based distinctions and presumptions of expertise based solely or principally on years of professional experience. The findings of this thesis open up new areas of research where the identification and ranking of legal experts based on readily identifiable and measurable cognitive differences may now be approached with greater confidence and specificity.

A *The Research Question*

The research question addressed in this thesis is:

> In what readily identifiable and measurable ways do legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts?

The significance of this question and the issues it raises lies in the centrality of cognitive skills to a lawyer's training and professional work. The idea that lawyers think differently from other people is ingrained in modern culture. The dimensions of these differences, however, have not been fully investigated in a scientific sense. Even less understood and less researched are the cognitive differences that exist between lawyers with different levels of legal expertise. This is the context in which the above research question fits, and where the findings of this thesis have their greatest theortical and pedagogical significance.

Not only does this thesis open up new areas for investigation, its findings have immediate and direct applicability in a range of situations. These include accreditation schemes within the legal industry, lawyer-selection choices by clients, legal training programs and law-school education initiatives. Moreover, the methodology used in this research project offers guidance to future researchers seeking both to reproduce the findings reported here and to extend the qualitative and quantitative framework developed for this study. No previous studies have provided an equivalent foundation based on an ability to distinguish between different levels of expertise within a single field of legal specialisation.

As to the form of the research question chosen for this thesis, the focus on readily identifiable and measurable differences ensured that the findings of this study would be of practical use to future researchers as well as to others with an interest in lawyer accreditation schemes, the development of legal talent within law firms, enhancing information available to consumers of legal services, and legal education particularly at intermediate and advanced levels. While theoretical insights about cognitive processes were important, the ultimate aim was to generate a list of identifiers, traits and indicators that could be easily and reliably used by a variety of interested parties to distinguish between different levels of legal experts within a single field of legal specialisation.

Restricting the research question to information-limited and time-constrained contexts reflected the empirical basis of the study which was informed, in part, by the need to establish a benchmark for future testing that would likely take place in similar contexts, for example, by way of written or oral examinations. Assessing legal expertise in a purely naturalistic setting was not considered feasible given the methodological tools available. Nor would it have been as useful for future researchers and other interested parties given their likely reliance on carefully selected case studies and examination-like test environments.

The chosen area of legal specialisation in this study was competition law. More specifically, it was Australian and New Zealand competition law with a specific focus on the merger clearance procedures that exist under the legislated regimes of both countries as well as in most other competition law jurisdictions around the world. The advantage of choosing competition law for this thesis was twofold. First, it is a distinct

and recognised field of legal specialisation such that practitioners can be easily identified. Second, the researcher has been both a legal and economic specialist in this area of law for over 20 years. This facilitated the efficient identification of technical legal issues and comports with previous studies along similar lines which have been conducted by researchers with extensive experience in immigration law, social security disability law and family law.

This choice of legal specialisation required a detailed analysis of competition law issues during the testing process. However, the findings of the thesis are general in nature and can be applied to a variety of other specialist fields of law. Since all participants had essentially the same technical legal knowledge, such knowledge was not in or of itself a distinguishing factor when comparing their test performances. The results from these tests can therefore be compared directly to the higher-level reasoning strategies of legal practitioners in other specialist areas of law.

In terms of methodology, this thesis adopts the well-established techniques of think-aloud problem solving and verbal protocol analysis. These investigatory techniques have been used successfully in previous studies of legal thinking skills. Based on the theoretical work of cognitive psychologists, this approach to eliciting and analysing ecologically valid verbal data was considered ideally suited to this research project which aims to collect empirical data on how lawyers actually think. Moreover, by maximising the number of study participants and the assessment tasks they undertook, this study identifies statistically-significant differences between the cognitive skills of participants with different levels of specialist expertise.

Lastly, this thesis' response to the research question was aligned with the traditional categories of expertise development, namely, the progression from apprentice to journeymen and from expert to master. By design, no participant in this study had less than five years specialist experience in competition law. Accordingly, it was not necessary to consider the lower levels of novice, initiate or layman. Further background to these traditional categories of expertise development is provided in the next chapter.

B *Summary of Findings*

The findings of this thesis are based on an analysis of 73 case assessments undertaken by 20 legal specialists, all of whom were required to think-aloud as they analysed and assessed legal risk in each case. These findings can be broken down into qualitative identifiers, performance traits and uses of intuition which together can be used to distinguish legal specialists with different levels of expertise. The following table summarises a number of these expertise-related identifiers and cognitive attributes.

TABLE 1.1 – Summary of Expertise-Related Identifiers, Performance Traits and Intuitive Responses of Study Participants

	1. IDENTIFIERS	2. PERFORMANCE TRAITS	3. INTUITIVE RESPONSES
MASTER	• **Correct legal-risk assessments based on substantive analyses** • Unlaboured reasoning	• **Identifies key issues easily, and very quickly** • Avoids irrelevant and low-quality issues • Uses appropriate and specific analogies • Ignores/glosses-over short-cuts that minimise synthesis • Synthesises issues efficiently and effectively	• **Heavy and effective reliance on intuition** • Extensive and effective self-monitoring • **Avoids deliberative doubt**
EXPERT	• Unlaboured reasoning • Inconclusive assessments • **Refusals to give an assessment when insufficient information available**	• **Identifies key issues easily** • Avoids irrelevant and low-quality issues • Uses appropriate and specific analogies • Ignores/glosses-over short-cuts that minimise synthesis • Synthesises issues efficiently and effectively	• **Integrated intuition and reasoning** • Extensive and effective self-monitoring • **Limits deliberative doubt**
JOURNEYMAN	• **Correct legal-risk assessments based on superficial analyses** • Unlaboured reasoning • Inconclusive assessments • Incorrect legal-risk assessments	• **Issue identification is laboured**. • **Identifies short-cuts that minimise synthesis** • Identifies irrelevant and low-quality issues • Synthesis is inefficient and limited/superficial	• **Substitutes guessing for intuition** • Limited and ineffective self-monitoring • **Tendency to 'blurt out' premature responses**
APPRENTICE	• **Laboured reasoning** • **Inconclusive assessments** • Incorrect legal-risk assessments	• **Issue identification is laboured, and extensive** • **Uses vague and inappropriate analogies** • Identifies irrelevant and low-quality issues • Synthesis is inefficient and limited/superficial	• **Heavy reliance on analytical reasoning** • Limited self-monitoring • **Overwhelmed by deliberative doubt**

These findings indicate that under test conditions such as those utilised in this study, master-level legal specialists can be identified by their correct substantive analysis of legal issues, the ease and speed with which they identify key issues, and their heavy reliance on intuition in assessing legal risk. Experts can similarly be identified by their refusals to provide concluded views when information and time are limited, their ability to identify key issues quickly, and their effective integration of intuition and analytical reasoning.

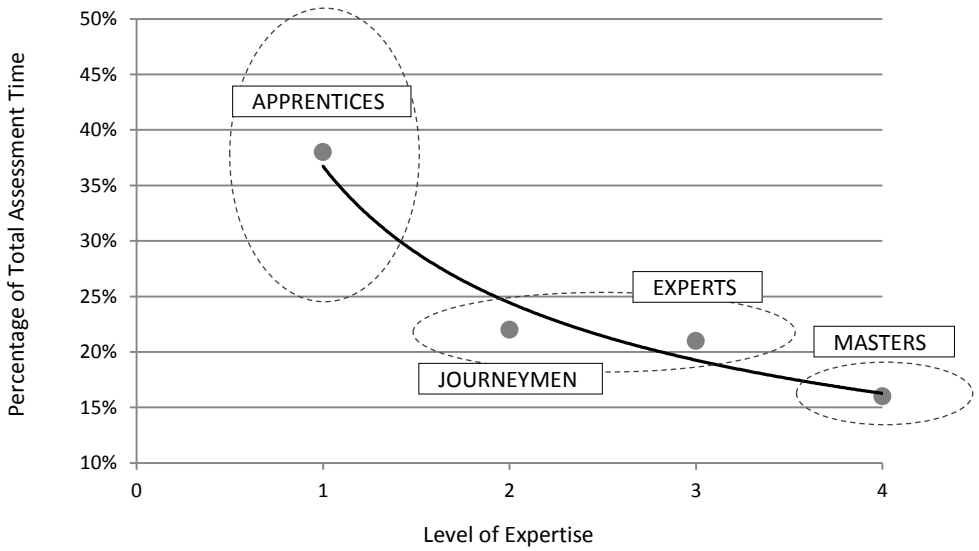
In terms of lower-level legal specialists, journeymen are identifiable by their reliance on superficial reasoning to assess legal risk, their laboured identification of issues, and their tendency to rely on guessing rather than their intuition and to 'blurt out' their responses prematurely. Apprentices are identifiable by their laboured reasoning and extensive search for relevant issues and analogies, their heavy reliance on analytical reasoning, and the significant influence of self-doubt on their deliberations.

A more detailed discussion of these and other expertise-related identifiers, performance traits and intuitive responses is provided in Chapter 6.

1 *Quantitative Differences*

The findings summarised in Table 1.1 are based, in large part, on statistically-significant quantitative differences between the legal-risk assessment performances of study participants. Some of these differences can be expressed as numeric ratios and charted to illustrate relationships and associations between particular cognitive performance characteristics and different levels of specialist legal expertise. The following chart, which depicts time spent retrieving information from long-term memory (denoted here as identifying issues) as a proportion of total assessment time, provides an example of a relevant expertise-based association.

CHART 1.1 – Average Proportion of Total Time Spent Identifying Issues

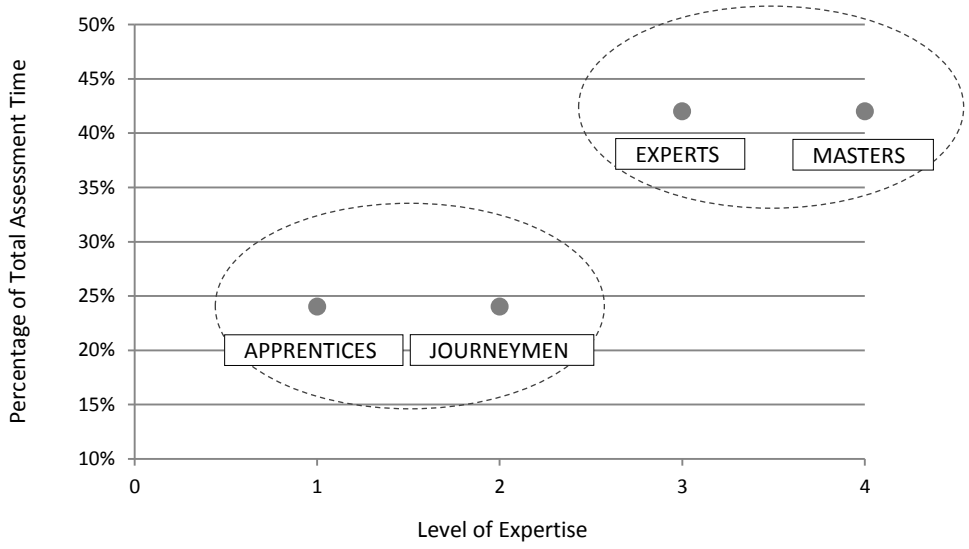


This chart indicates that as expertise increases, legal specialists spend less time retrieving information from their long-term memories. This is ostensibly because they

have a better sense of which issues are most important compared to less expert legal specialists who have a less developed sense of issue priorities and relevance, as well as less efficiently organised mental schema. As between journeymen and experts, however, no statistically significant difference was identified on this measure. The fact that master-level participants spent less time than experts on this task suggests that cognitive performance in this area continues to improve even amongst highly-experienced lawyers.

A similar chart can be generated from data relating to the proportion of total assessment time participants spent drawing inferences when assessing legal risk in the test-cases. The chart below shows that higher-level (more expert) participants in this study spent substantially more time drawing inferences (denoted as synthesis) compared to lower-level participants. However, amongst experts and masters no further distinction was evident. There was also no statistically significant difference between apprentices and journeymen on this measure.

CHART 1.2 – Average Proportion of Total Time Spent Engaging in Synthesis



Cognitive load statistics were also recorded during the legal risk-assessment process by measuring participants' rates of verbalisation as they engaged in different cognitive tasks. These data indicate that apprentices and journeymen exert greater effort than experts and masters while retrieving information from their long term memories and when drawing inferences. In addition, whereas experts and masters exert the same cognitive effort whether identifying or synthesising issues, apprentices and journeymen

tend to experience greater cognitive load when identifying issues and appear to find synthesis significantly easier in relative terms, ostensibly because the inferences they draw are more superficial.

2 *Previous Studies*

This thesis relied on a variety of previous studies to guide its methodological design. It also confirmed the findings of researchers who have investigated the cognitive behaviours and abilities of experts in general and lawyers in particular. These confirmations were, however, largely incidental to the primary objective of developing a response to the thesis' research question regarding readily identifiable and measurable expertise-related cognitive differences. In terms of cognitive traits previously identified by researchers of general expertise, the following confirmatory results were recorded:

- Higher-level (ie more expert) legal specialists in this study used more effective and more efficient risk assessment strategies, with masters being able to correctly evaluate risk based on substantive analysis notwithstanding the information and time-constraints that existed. They were also the only participants able to recall 'instant' solutions (that is, solutions that appeared to avoid intermediate steps) on this basis.

- There was some evidence that the higher-level specialists in this study were less flexible in their methods and more dependent on contextual information, which may explain the refusals by two expert/master-level participants to provide any assessment because of insufficient background information.

- The atypical case used in this study revealed differences between higher-level participants who either provided an instant solution based on their expert knowledge or successfully undertook a substantive and conceptually difficult assessment, and lower-level participants who were either attracted to analytical short cuts (which resulted in superficial analyses) or struggled to complete their assessments.

- The responses of some higher-level participants suggested that they had glossed-over unimportant details in the test-case documents, while their more effective self-monitoring skills could explain why they verbalised fewer comprehension and analytical errors than lower-level participants.

Regarding those researchers who previously focused on the cognitive analysis of legal expertise, further confirmations were apparent from the data collected during this study. These confirmations include:

- Higher-level participants in this study were faster at diagnosing relevant legal issues, at least in the sense of more quickly identifying the key factual information in the test-case documents. (The area of law itself was known at the start of the tests, so there was no testing of whether or not participants could identify the applicable legal rules.)

- The better performance of higher-level participants was likely to have been explained by their having more effective mental schemas for assessing legal risk in the context of the test-cases used in this study.

- Higher-level participants also had a clearer and more specific idea of what further information they required – and they had more effective strategies for how to get that information. Lower-level experts tended to be more vague in these areas.

- Lower-level participants generally exhibited less precise and more generalised use and recall of information. This was particularly evident in instances of laboured reasoning and in lower-level participants' greater focus on retrieving information which was generally of a lower quality than that relied upon by higher-level participants.

Lastly, the results of this study suggest that some previous researchers have relied on questionable assumptions regarding the expertise levels of their study participants. In studies concerning generic legal skills such as reading court cases, this is unlikely to have been a significant issue. However, for those studies which sought to compare the

cognitive performances of legal experts and novices on more substantive tasks, there may have been material – yet unacknowledged – variation in the cognitive abilities of the presumed experts. Whether or not this warrants a reassessment of those studies, future studies are likely to benefit from the findings of this thesis, which provides a list of readily identifiable and measurable differences in the cognitive performances of legal specialists from five years through to more than 30 years of specialist experience.

C *Chapter Outline*

This thesis is divided into six further chapters.

Chapter 2 identifies a lack of scholarly literature concerned with the cognitive skills of legal specialists and describes how the few empirical studies that have taken place in this area have been restricted to expert-novice comparisons rather than the higher-level comparisons of interest to this thesis. This review summarises the nature of this gap in the literature while at the same time identifying the methodologies likely to be most useful for further research in the area. This includes a discussion of the conceptual framework of progressive expertise, which has to date largely been overlooked in favour of knowledge-based expertise frameworks as used by cognitive psychologists investigating expert thinking in medicine, amongst other professions.

Chapter 3 describes the process by which the methodology used in this study was developed having regard to previous research amongst legal scholars and researchers in other fields who have used think-aloud verbal protocol analysis to investigate cognitive traits associated with expert performance. It describes the five measures used to rank participants in this study according to their levels of apparent or likely expertise. It also explains the procedures used to elicit and record the concurrent verbalisations of participants as they assessed legal risk in the provided test-cases. The chapter concludes with a discussion of the exploratory nature of the analytical approach adopted in this thesis to describe readily identifiable and measurable differences in the cognitive performances of study participants with different levels of specialist legal expertise.

Chapter 4 explains how 20 competition law specialists spread across the major business centres of Australia were selected and then ranked according to their levels of likely expertise using the methodological framework developed in Chapter 3. This

explanation includes detailed accounts of the processes by which individual lawyers were assessed having regard to their work experience and their performance in assessing the test cases. The chapter includes a comparative review of rankings from independent research publications that purport to identify leading lawyers in various specialist fields in different jurisdictions.

Chapter 5 describes how the results of participants' test performances were assessed and categorised according to an analytical framework that sought to highlight readily identifiable and measurable differences between higher and lower-ranked participants. This includes an overview of the data collected and the application of the methodologies described in Chapter 3. The substantive discussion in this chapter centres on the qualitative and quantitative analysis of participants' verbal transcripts and the differences that became evident during the comparative process. The chapter concludes with a summary of observed behavioural and quantitative differences.

Chapter 6 records a deeper analysis of the results observed in the Chapter 5. It begins by focusing on associations between levels of expertise and time spent identifying issues and drawing inferences. It then analyses verbalisation rates for each category of participant and for specific cognitive tasks. After this, a further analytical approach is considered to investigate possible reasons why participants behaved in the ways that were observed. This is followed by a discussion of several factors that may have influenced the results of the study. The final section of the chapter summarises the findings of the study for the purpose of providing a direct response to the research question in both tabular and written forms. These findings include a list of the behavioural identifiers, performance traits and cognitive indicators that distinguished the different thinking approaches and abilities of the apprentice, journeyman, expert and master-level legal specialists in this study.

Chapter 7 reviews the previous research on which this study was based and identifies areas where this thesis either confirms or raises questions about the methodology and findings of previous studies that have investigated the cognitive aspects of legal expertise. It also discusses the limitations of the present study in terms of design and execution. It then considers the areas of future research facilitated by or likely to benefit from the methodological approach adopted here and the responses provided to the original research question. These areas of future research extend across the

disciplines of cognitive psychology, lawyer accreditation schemes, assessments by consumers of legal services, talent management within law firms, and legal education.

II LITERATURE REVIEW

This chapter begins with a review of the scholarly literature concerned with legal thinking. It then considers various methodologies and conceptual frameworks that have been used to investigate and analyse the cognitive characteristics of legal expertise. The discussion concludes by confirming the need for further research to provide a response to this thesis' research question insofar as there does not presently exist a list of identifiers, traits and indicators that can be used to measure specialist legal expertise beyond the knowledge-based, novice-expert dichotomy.

Part A of this chapter explores both scholarly and popular interest in the phrase 'to think like a lawyer.' This includes a review of existing research involving legal experts which to date has been limited to the comparison of novice and expert differences, but not the cognitive abilities of higher-level legal specialists within the same field of law. Given the focus of the present study on how specialist lawyers with different levels of expertise think, scholarly contributions to the debate over what it means to think like a lawyer are identified as lacking sufficient empirical data to ground a compelling response to this study's research question.

Part B considers previous empirical research into the cognitive traits of lawyers. Researchers in this area have combined theoretical research on legal thinking skills with methodologies developed by cognitive psychologists to trace the mental processes that distinguish the mental abilities of experienced lawyers. However, these studies have predominantly been undertaken by scholars outside the legal academy and have been restricted to generic legal skills, such as the effective reading of court cases. While suggestive of a methodological path towards answering this study's research question, no studies in this area offer an adequate response.

Part C examines a different methodological approach adopted by cognitive psychologists who have studied expertise amongst medical practitioners. This approach acknowledges the discrete nature of medical specialisation. Scholars in this field have developed analytical schemes and frameworks to distinguish between different levels of expertise within and amongst medical specialists. However, the work of these researchers has been confined to studying knowledge-based differences,

where higher-level experts are distinguished by their possession of specialist medical knowledge. This is a different conceptual approach from that required to measure cognitive differences between legal specialists with the same technical legal knowledge, but different levels of expertise.

Part D draws together the work of legal scholars who have undertaken or foreseen the potential for empirical research into the cognitive traits of legal specialists with different levels of expertise. It identifies the influence of previous research from the area of medicine which has resulted in a refocusing of scholarly attention on how legal specialists actually think, even though this refocusing has largely been limited to knowledge-based inquiries. This discussion confirms that the cognitive differences noted in this context are those attributable to differences in specialist legal knowledge rather than differences in levels of expertise within the same specialist field of law.

Part E discusses an alternative conceptual framework better suited to identifying cognitive differences between legal specialists with different levels of expertise within the same field of legal specialisation. This framework draws on the traditional progressive categorisation of expertise, which both presumes a basic common knowledge across all levels along the specialist expertise continuum and provides descriptions of an individual's level of expertise as they move through different stages of development.

Part F describes an intra-specialist conceptual framework within which lawyers with the same specialist legal knowledge may be compared according to different cognitive skills associated with their different levels of specialist legal expertise. This discussion confirms the efficacy of progressive expertise frameworks in general, and highlights the relevance of the traditional model as both a workable and appropriate methodological construct for the current proposed research into the identification and measurement of the expertise-related cognitive traits of legal specialists.

The chapter concludes that the research question for this study cannot be answered by previous research into how lawyers think. This is not because the methodologies or techniques that have been developed are not appropriate or are unlikely to be effective in an intra-specialist comparative context. Rather, it is because no previous studies have sought to identify and analyse the cognitive differences that are of interest here. Those studies have been predominantly concerned with knowledge-based comparisons

rather than with the progressive conceptualisation of specialist legal expertise. Given this gap in the existing scholarly literature, the latter approach was adopted to answer this thesis' research question.

A *To Think Like a Lawyer*

The ways in which lawyers solve legal problems have long been of interest to researchers of expertise and expert performance, perhaps in part because of the perceived uniqueness of what it means 'to think like a lawyer.' Yet empirical studies have not been as numerous nor as extensive as those involving chess players and medical doctors, to name just two of the hundreds of fields of human endeavour in which experts have been identified and their cognitive abilities subjected to investigation. As a consequence, the empirical study of lawyering expertise remains a promising though largely untapped area of research, especially at the higher levels of legal specialists.

One of the most important contributions to the study of legal expertise from scholars involved in wisdom and expertise research is the application of methodologies that enable the tracing of cognitive processes. Amongst these methodologies, one of the most popular and best understood is think-aloud verbal protocol analysis.[1] This methodology, which involves participants engaging in 'think-aloud' problem-solving tasks, has significant advantages over interviews, surveys and other data gathering techniques because it permits direct investigation of how individuals actually think.

[1] The popularity of this methodology, particularly amongst the leading researchers in wisdom studies, is well documented. See Monika Ardelt, 'Wisdom as Expert Knowledge System: A Critical Review of a Contemporary Operationalization of an Ancient Concept' (2004) 47 *Human Development* 257. The so-called Berlin group of wisdom researchers, whom Ardelt describes as having undertaken 'the most extensive and systematic empirical work on wisdom in the field to date' (Ibid 258) and whom Robert Sternberg calls 'world leaders in the study of wisdom' (Robert J Sternberg, 'Words to the Wise About Wisdom: A Commentary of Ardelt's Critique of Baltes' (2004) 47 *Human Development* 286, 286), have consistently embraced think-aloud verbal protocol analysis. See, as examples: PB Baltes and UM Staudinger, 'Wisdom: A Metaheuristic (Pragmatic) to Orchestrate Mind and Virtue Toward Excellence' (2000) 55 *American Psychologist* 122; PB Baltes, UM Staudinger, A Maercker and J Smith, 'People Nominated as Wise: A Comparative Study of Wisdom-Related Knowledge' (1995) 10 *Psychology and Aging* 155; J Smith and PB Baltes, 'Wisdom-Related Knowledge: Age/Cohort Differences in Response to Life-Planning Problems,' (1990) 26 *Developmental Psychology* 494; Staudinger, UM, J Smith and PB Baltes, 'Wisdom-Related Knowledge in a Life Review Task: Age Differences and the Role of Professional Specialization' (1992) 7 *Psychology and Aging* 271. For a summary of hundreds of studies that have used this methodology in the contexts of various professions see K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006). See also the recent review of almost 2,000 think-aloud studies in Mark C Fox, K Anders Ericsson and Ryan Best, 'Do Procedures for Verbal Reporting of Thinking Have to Be Reactive? A Meta-Analysis and Recommendations for Best Reporting Methods,' (2011) 137(2) *Psychological Bulletin* 316.

Accordingly, while there is a relative lack of empirical research on how lawyers actually think, the tools for undertaking such investigations are both readily available and well understood. Moreover, the few cognitive studies that have involved lawyers have consistently demonstrated the feasibility of scholarly inquiry of this kind.[2] These studies have not, however, significantly advanced our understanding of how legal practitioners at the highest levels in specialist fields of law think. This is largely because previous researchers have typically compared novices with presumed experts in tasks involving generic or general legal skills. Only a few have compared intra-specialist cognitive differences within the same field of law.

The application of think-aloud verbal protocol analysis to lawyers can be traced back to the mid-1970s,[3] although use of this methodology in other knowledge domains began as early as the 1920s.[4] One reason why cognitive researchers began to focus their attention on lawyers was because of a long held belief that 'to think like a lawyer' involved thinking differently from other people.[5] This view has persisted within

---

[2] See, as examples, Mary A Lundeberg, 'Metacognitive Aspects of Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407; Dorothy H Deegan, 'Exploring Individual Differences Among Novices Reading in a Specific Domain: The Case of Law' (1995) 30(2) *Reading Research Quarterly* 154; Fernando Colon-Navarro, 'Thinking Like a Lawyer: Expert-Novice Differences in Simulated Client Interviews' (1997) 21 *The Journal of the Legal Profession* 107; Ian Weinstein, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1; James F Stratman, 'When Law Students Read Cases: Exploring Relations Between Professional Legal Reasoning Roles and Problem Detection' (2002) 34(1) *Discourse Processses* 57; James F Stratman, 'How Legal Analysts Negotiate Indeterminacy of Meaning in Common Law Rules: Toward a Synthesis of Linguistic and Cognitive Approaches to Investigation' (2004) 24 *Language & Communication* 23; Laurel Currie Oates, 'Leveling the Playing Field: Helping Students Succeed by Helping Them Learn to Read as Expert Lawyers' (2006) 80 *St John's Law Review* 227; Leah M Christensen, 'Legal Reading and Success in Law School: An Empirical Study' (2006-2007) 30 *Seattle University Law Review* 603; Leah M Christensen, 'The Paradox of Legal Expertise: A Study of Experts and Novices Reading the Law' (2008) *Brigham Young University Education and Law Journal* 53; Fleurie Nievelstein, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Expertise-Related Differences in Conceptual and Ontological Knowledge in the Legal Domain' (2008) 20(6) *European Journal of Cognitive Psychology* 1043; Fleurie Nievelstein, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Effects of Conceptual Knowledge and Availability of Information Sources on Law Student's Legal Reasoning' (2010) 38 *Instructional Science* 23; Fleurie Nievelstein, Tamara van Gog, Gijs van Dijk and Henny P A Boshuizen, 'Instructional Support for Novice Law Students: Reducing Search Processes and Explaining Concepts in Cases' (2011) 25 *Applied Cognitive Psychology* 408.

[3] H F M Crombag, J L De Wijkerslooth and E H Van Tuyl Van Serooskerken, 'On Solving Legal Problems' (1975-1976) 27 *Journal of Legal Education* 168.

[4] See, for example, J B Watson, 'Is Thinking Merely the Action of Language Mechanisms?' (1920) 11 *British Journal of Psychology* 87; and, K Duncker, 'A Qualitative (Experimental and Theoretical) Study of Productive Thinking (Solving Comprehensible Problems)' (1926) 33 *Pedagogical Seminary* 642.

[5] One of the earliest recorded instances of the phrase 'to think like a lawyer' (written within quotation marks) was in the first edition of the *Journal of Legal Education* in 1948, in a report on the proceedings of the first National Law Student Conference in the United States. John De J Pemberton, 'National Law Student Conference: The Conference Report' (1948-49) 1 *Journal of Legal Education* 73, 90. See also Clarence Morris, *How Lawyers Think* (Harvard University Press, 1937).

popular culture and within the legal academy, and ostensibly explains at least in part the marketing of self-help books based on this premise,[6] as well as the numerous scholarly references to the 1973 Hollywood movie *The Paper Chase*, in which Professor Charles Kingsford famously pronounced to his class of first year law students, 'You teach yourselves the law. I train your minds. You come in here with a skull full of mush, and if you survive, you'll leave thinking like a lawyer.'[7]

For a significant number of legal academics, Professor Kingsford's speech has become something of an intellectual touchstone, arguably because of the paucity of empirical studies on legal thinking skills.[8] For some scholars, just as it was for the fictitious Professor, the existence of a specific legal way of thinking is beyond question. This view permeates the myriad formalised methods of legal reasoning taught to law students during the early years of law school. Such methods typically have their own acronyms, such as IRAC, HIRAC, IREAC, MIRAT and CREAC,[9] and their central

---

[6] What it might mean 'to think like a lawyer' has today become a key theme in a number of books aimed at a general readership promising to impart the 'secrets' of legal thinking skills to lay people – as if the ability to think like a lawyer is a special attribute and an advantage only experienced by those people who have attended law schools. See as examples, Robert J Dudley, *Think Like a Lawyer: How to Get What You Want by Using Advocacy Skills* (Nelson-Hall, 1980); Kenneth J Vandevelde, *Thinking Like a Lawyer: An Introduction to Legal Reasoning* (Westview Press, 1996) – The marketing spiel on this book states: 'This is not a book about the content of the law; it is about a well-developed and valuable way of thinking that can be applied to many fields'; Ruth Ann McKinney, *Reading Like A Lawyer: Time-Saving Strategies for Reading Like an Expert* (Carolina Academic Press, 2nd ed, 2012); Larry Kahn, *Think Like a Lawyer: Negotiating a New Car Purchase* (Amazon Digital Services, 2012). The opportunity should not be missed to note also that blogs and other easily accessible sources of information that inform popular opinion are not just a repository of hilarious lawyer jokes, but also of observations like this: 'In case you were wondering, here is a formal definition of *to think like a lawyer*: verb, 1) to display a neurotic tendency to hostile, oppositional behavior; and/or 2) to obsess over unnecessary, pointless detail.' The People's Therapist, 'Oversold' (25 April 2012) <http://thepeoplestherapist.com/2012/04/25/oversold/ >.

[7] *The Paper Chase* (Directed by James Bridges, Twentieth Century Fox, 1973). John Jay Osborne's book on which the movie was based, also titled *The Paper Chase*, was written while the author attended Harvard Law School. During a visit to his alma mater in October 2012, Osborne said he was motivated to write the book after sensing that the law school 'glorified their teachers over their students.' *'"The Paper Chase" at 40: Law School audience reflects on iconic film about earning degree,'* (Harvard Gazette, 2 October 2012) <http://news.harvard.edu/gazette/story/2012/10/the-paper-chase-at-40/>.

[8] As many as half of all the articles reviewed by the researcher in which a legal scholar discussed teaching students how to think like lawyers, cited – and often quoted from – Professor Kingsfield's fictional speech.

[9] For an introduction to what these acronyms mean and where more information about them can be found see Nick James, 'Logical, Critical and Creative: Teaching 'Thinking Skills' to Law Students' (2012) 12(1) *QUT Law & Justice* 66. As to the above acronyms: IRAC = Issue, Rule/Law, Application, Conclusion; HIRAC = Heading, Issue, Rule, Application, Conclusion; IREAC = Issue, Rule, Explanation of rule, Application, Conclusion; MIRAT = Material facts, Issues, Rules, Arguments, Tentative conclusion; and, CREAC = Conclusion, Rule, Explanation of rule, Application of rule, Conclusion. See further: Jeffery Metzler, 'The Importance of IRAC and Legal Writing' (2002-2003) 60 *University of Detroit Mercy Law Review* 501; Soma Kedia,'Redirecting the Scope of First-Year Writing Courses: Towards a New Paradigm of Teaching Legal Writing' (2009-2010) 87 *University of Detroit Mercy Law Review* 147; Michelle Sanson, Thalia Anthony and David Worswick, *Connecting with Law* (Oxford

purpose is to assist law students identify legal issues and present their solutions to legal problems in a lawyerly manner.[10] In other words, they replicate or mechanize what it means to think like a lawyer. In this regard, they reflect how law professors believe lawyers think or – cast in a normative pedagogical sense – should think.[11]

The continuing refinement of such methods has been accompanied by a growing number of reports and studies commissioned globally to clarify what law schools should be teaching their students, and more specifically what kinds of pedagogical outcomes they should be aiming to achieve. For instance, the Australian Learning and Teaching Council's Threshold Learning Outcome 3 ('TLO3') for the Bachelor of Laws degree titled 'Thinking Skills' was formulated to align with studies, reports and recommendations published in a range of other jurisdictions, including Britain, the United States, Canada and Scotland.[12] Its accompanying Good Practice Guide to

---

University Press, 2nd ed, 2010); ANU Academic Skills and Learning Centre, 'Legal Reasoning and HIRAC' (2010); Mark E Wojcik, 'Add an E to Your IRAC' (2006-2007) 35 *Student Lawyer* 26; John Wade,'Meet MIRAT: Legal Reasoning Fragmented into Learnable Chunks' (1990-1991) 2 *Legal Education Review* 283; David S Romantz and Kathleen Elliot Vinson, *Legal Analysis: The Fundamental Skill* (Carolina Academic Press, 2nd Ed, 2009).

[10] With reference to the above acronyms, James wrote that 'Students are taught how, when presented with a set of facts in the form of a tutorial problem or an exam question, they should identify the legal issues and, considering each issue carefully and logically, apply the relevant legal rules to the facts in order to reach a rational and convincing conclusion about the legal consequences of the particular situation.' James, above n 9, 76.

[11] Although it should also be noted that some academics argue against the use of these formalistic approaches in legal practice claiming they are unlikely to facilitate the flexible, tailored analyses required in real world settings. See, for instance, Bryan A Garner, 'Although IRAC Works for Exams, Avoid it in Practice' (2004-2005) 33 *Student Lawyer* 10. James further states that 'The prevailing view in Australia appears to be that formalistic techniques such as IRAC are useful for students new to the study of law, but as they progress through their legal studies the "scaffolding" offered by the step by step techniques should recede into the background in favour of a greater emphasis upon "flow" in the student's reasoning and consequent improvements in subtlety and persuasiveness.' Ibid.

[12] The authors of these TLOs summarized the various – and consistently general – ways in which researchers and advisory bodies in other jurisdictions have described what they consider the legal thinking skills that law school graduates should be expected to have developed:

> The United Kingdom QAA Subject Benchmark Statement for Law stresses the need for graduates to be able to 'make a critical judgment of the merits of particular arguments,' 'critical analysis' being 'recognised as a key attribute of graduates.' Also, the QAA identified that graduates should have an ability to 'present and make a reasoned choice between alternative solutions.' This is similar to the United Kingdom Joint Statement of the Law Society and the General Council of the Bar's requirement that graduates should be able to 'recognise potential alternative conclusions for particular situations, and provide supporting reasons for them.' The MacCrate Report in the United States recognised that graduates should 'be familiar with the skills and concepts involved in identifying and formulating legal issues.' The Task Force on the Canadian Common Law Degree recommended that entrants to a Canadian bar admission program be required to have demonstrated skills in solving legal problems that included the ability to 'a. identify relevant facts; b. identify legal, practical, and policy issues and conduct the necessary research arising from those issues; c. analyse the results of research; d. apply the law to the facts; and e. identify and evaluate the appropriateness of alternatives for resolution of the issue or dispute.' The Scottish Accreditation Guidelines require that graduates have a basic competence in 'apply[ing] knowledge and analysis... creatively to complex situations in order to

teaching TLO3 includes a review of relevant literature such as text books and articles dedicated to explaining various reasoning strategies using the above reasoning methods and related strategies.[13] As to the elements of TLO3 itself, which is one of six TLOs developed under the Learning and Teaching Academic Standards Project, these recommend that law graduates should be able to '(a) identify and articulate legal issues, (b) apply legal reasoning and research to generate appropriate responses to legal issues, (c) engage in critical analysis and make a reasoned choice amongst alternatives, and (d) think creatively in approaching legal issues and generating appropriate responses.'[14]

Kift, Israel and Field describe how they identified these four essential elements of legal thinking after undertaking national consultations with industry and academic representatives, and following an investigation of international trends amongst legal educators.[15] They note the universal emphasis on graduates being capable of thinking 'analytically and creatively in approaching and generating solutions to legal issues,' independent of any specific legal knowledge in a substantive sense.[16] This approach comports with the authors' characterization of legal education as a continuum, just as their Canadian counterparts had focused on 'foundational competencies necessary for the practice of law' while acknowledging the important role of the legal profession in providing additional inputs along the path towards formal licensing.[17]

The key challenge for theoreticians has been how to define (and not merely describe) what constitutes lawyerly thinking, a concept which is viewed by many scholars as critical to the mission of law schools.[18] Some authors have written about what they

---

provide arguable solutions to concrete problems by presenting a range of viable options from a set of facts and law' and also that they '[t]hink critically and make critical judgments on the relative and absolute merits of particular arguments and solutions.'

Sally Kift, Mark Israel and Rachael Field, Learning and Teaching Academic Standards Project: Bachelor of Laws Learning and Teaching Academic Standards Statement, December 2010 (Australian Learning and Teaching Council, 2010) 17 (references omitted). Also see James, above n 9.

[13] Nick James, *Good Practice Guide (Bachelor of Laws): Thinking Skills (Threshold Learning Outcome 3)* (Australian Learning and Teaching Council, 2011).

[14] Ibid 1.

[15] Kift, Israel and Field, above n 12.

[16] Ibid 8.

[17] Task Force on the Canadian Common Law Degree, *Final Report* (2009) 23.

[18] '[T]he critical first step to understanding the purpose of law school is to define what it means to "think like a lawyer." Although this is an oft-stated justification for law school, no consensus exists about what "thinking like a lawyer" means. Furthermore, the concept of thinking like a lawyer is loaded with notions, accumulated over time, about what a lawyer does. We need to re-examine and clearly define what it means to "think like a lawyer" before we can begin to assess lawyers' practices.' Bethany Rubin Henderson, 'Asking the Lost Question: What is the Purpose of Law School?' (2003) 53(1) *Journal of*

perceive to be a lack of agreement on 'a detailed conception of what "thinking like a lawyer" means,'[19] while others have called it 'the hoary old saw' of legal education.[20] These scholars claim that notwithstanding the popularity of this phrase and numerous articles and books on the subject, the definitions that have been offered are 'unfulfilling' and clearly show the need for a greater focus on the 'cognitive components of the skill that has become a central theme of legal education.'[21]

According to Schauer, the author of a recently reprinted text-book on thinking like a lawyer,[22] the manner in which law school professors typically escape confronting this definitional void is particularly revealing. He has observed how his fellow law-school deans, when asked by first year law students about the value of a legal education, begin to realise that 'we claim that we are teaching students how to think like lawyers, but we never really tell them what it is to think like lawyers; we assume that they will absorb this by osmosis, and we assume it the same way that people have assumed it for decades or generations.'[23] Which is, as Schauer also notes, essentially the same approach adopted by Professor Kingsfield who similarly avoided providing any definition.[24]

Other legal academics have argued that not only is there such a thing as 'thinking like a lawyer,' but that depending on what kind of law is involved different types of legal thinking are required. Okamoto, for instance, has argued that by exposing law students to the mathematics of business and finance, their ability as transactional lawyers can be greatly improved such that 'we will have taught them how to begin to think like a "deal lawyer" – something very different from what we have traditionally meant by "thinking

---

Legal Education 48, 57. See also Larry O Gantt, 'Deconstructing Thinking Like a Lawyer: Analyzing the Cognitive Components of the Analytical Mind' (2006-2007) 29 *Campbell Law Review* 413.

[19] Gantt, above n 18, 413.

[20] Nancy B Rapoport, 'Is "Thinking Like a Lawyer" Really What We Want to Teach?' (2002) 1 *Journal of the Association of Legal Writing Directors* 91, 104.

[21] Gantt, above n 18, 415.

[22] Frederick Schauer, *Thinking Like a Lawyer: A New Introduction to Legal Reasoning* (Harvard University Press, 2009, reprint 2012).

[23] Frederick Schauer, 'Do Lawyers Think, and If So, How?' presentation at an alumni luncheon on 30 April 2010 uploaded to Youtube.com by the University of Virginia School of Law on 4 May 2010 at http://www.youtube.com/watch?v=_l7FcYzvl7k accessed on 23 April 2012.

[24] Further support for the 'osmosis' theory of learning legal thinking skills can be found in the view expressed by James (whose opinion may well be shared by a large majority of law professors) that 'Legal reasoning is taught implicitly and constantly throughout a law student's legal studies in the sense that they are called upon to engage in legal problem solving in almost all their law subjects.' James, above n 13, 11.

like a lawyer".'[25] This is notwithstanding that, as noted, there is no clear consensus as to what that traditional meaning might be.

Okamato's contention conflicts with the claimed ability of law professors to teach 'the specialized type of thinking and problem solving *common to the legal profession*.'[26] His 'deal lawyer' thesis postulates that the yet-to-be defined concept of legal thinking does not exist in a singular form, but rather has a plurality of forms. For every specialist field of law, it may be argued, different legal thinking is required. Alternatively, whereas the pluralistic view is concerned with specialist legal practice post law-school, the standard view has more generic – even populist – aims. Whether the idea of specialist legal thinking therefore represents a progressive transformation of thinking skills, or merely a different perspective on the same theme, is unclear.

Still other scholars have questioned whether lawyers actually do think differently from other people. Stratman, for instance, queries 'whether legal thinking, legal reasoning, and legal argument comprise a unique kind of thinking, reasoning, and argument, that is, distinct from ordinary critical thinking, reasoning, and argument.'[27] But such views have not, on the whole, been expressed by legal academics. Rather they have been contributed by scholars from outside the legal academy, perhaps by researchers who may not appreciate the popularity of the belief amongst law professors in a distinct method of legal thinking that law schools are uniquely destined to teach.[28]

That said, a review of the above-listed elements of the TLO3 developed as the Australian benchmark for law graduate thinking skills reveal generalities both in terms of law (inasmuch as the listed attributes are non-specific and can relate to any area of

[25] Karl S Okamoto, 'Teaching Transactional Lawyering' (2009) 1 *Drexel Law Review* 69, 124.
[26] Media statement on the website of the Law School of the University of Virginia, where Professor Schauer teaches (Emphasis added). See *Professor's New Book Explores Legal Thinking* (30 April 2009) < http://www.law.virginia.edu/html/news/2009_spr/schauer_book.htm>.
[27] James F Stratman, 'The Emergence of Legal Composition as a Field of Inquiry: Evaluating the Prospects' (1990) 60 *Review of Educational Research* 2, 153, 161. Stratman cited a dozen contributors to this debate over whether lawyers do think differently from other professionals. Some of these arguments were being made as far back as the 1950s. See also Eric Amsel, Rosanna Langer and Lynn Loutzehiser, 'Do Lawyers Reason Differently From Psychologists? A Comparative Design for Studying Expertise' in R Sternberg and P Frensch (eds), *Complex Problem Solving: Principles and Mechanisms* (Lawrence Erlbaum Associates, 1991).
[28] 'Of all the verities of the faculty lounge, none is more plausible, or more widely believed, than the idea that law professors teach students how to think like lawyers.' David P Bryden, 'What Do Law Students Learn? A Pilot Study' (1984) 34 *Journal of Legal Education* 479, 479; 'The ability to "think like a lawyer" is frequently identified as one of the most important outcomes of the study of law.' James, above n 9, 68; 'The critical first step to understanding the purpose of law school is to define what it means to "think like a lawyer".' Henderson, above n 18, 57.

law) and in terms of describing standard forms of ordinary human reasoning. Being able to identify and articulate legal issues, for instance, may require special training to teach relevant legal knowledge. But the ability to engage in critical analysis and make reasoned choices amongst alternatives, or demonstrating the skills associated with creative thought, are those routinely taught to and possessed by professionals in many fields. The fact that TLO3 was developed by law professors could therefore suggest that at least some scholars within the legal academy also subscribe to the view that law schools essentially teach general thinking skills, albeit tailored or adapted to the practice of law.

A recent survey of legal practitioners revealed that they, too, are divided over what it means to think like a lawyer – which may be of even greater concern given that thinking like a lawyer is, at least on one view, their stock in trade. The definitions offered by survey respondents included statements that to think like a lawyer means: to think 'with both sides of the brain; precisely, rigidly, logically in relation to known rules;' 'Analyse things. Argue the alternative. Ask why;' 'having a jaundiced view of human nature;' 'find legal topics interesting. See the world in time limits and with reference to "what a judge might do";' engage in 'a lot of "worst case scenario" thinking;' 'be a good listener, lateral thinker, empathetic, analytical and persistent;' have a 'good ability to organise large amounts of information;' and ask, 'Is the client telling me everything?'[29]

It should be noted that the question put to survey respondents presumed a general or generic conception of legal thinking skills. Which is to say, no respondent was asked how a lawyer in a specialist field of law might think.[30] This, and a lack of guidance on what 'legal thinking' might mean (either from those conducting the survey or from the respondents' earlier law school experience), may explain the variety and disparity of responses recorded by the survey team.

[29] The survey question read as follows: 'In one or two sentences, what do you think it means to "think like a lawyer"? While the responses were predictably varied, it is possible to group them into four broad categories of thinking, ie, disciplined, solution-focused, realistic/pragmatic and ethical. Three quarters of all responses included comments falling within the first category, and half of these comments were accompanied by additional attributes falling within one or more of the other categories. Survey Monkey, *Thinking Like a Lawyer*, March 2012 (Copy of unpublished survey on file with the researcher).

[30] At the same time, none of the lawyers participating in this survey responded to this survey question by asking, 'What kind of lawyer?'

One of the researchers on this survey team had six years earlier reviewed the literature on what thinking like a lawyer entails.[31] In summing up the findings of her review, she admitted to not being a cognitive scientist herself, but merely a lawyer whose own views on this topic had 'been formed purely by an examination of "how" I think.'[32] Her article concluded that, 'it would be ideal if this topic were made the subject of proper scientific research as this could provide us as legal educators with great insights.'[33]

B *Cognitive Psychology and Lawyers*

Legal scholars began showing an interest in using methodologies from cognitive psychology, with a specific focus on think-aloud verbal protocol analysis, from the mid-1990s.[34] Their stated aim was to generate empirical data to inform their understanding of what it means to think like a lawyer.[35] But their contributions were pre-empted by researchers from outside the legal academy who were the first to use these methodologies to study the thinking processes of lawyers starting from the mid-1980s[36] although there was one earlier recorded but unpublished study undertaken in the mid-1970s.[37] It was not until the late 1990s that legal academicians began to undertake their own think-aloud verbal protocol studies on the subject of how lawyers think.

---

[31] Michelle Sanson, 'Thinking Like a Lawyer' (2006) *International Bar Association Conference Newsletter*.
[32] Ibid.
[33] Ibid footnote 28.
[34] The most influential theoretical discussion of these methodologies applied in a legal context was by law professor Gary Blasi, whose 1995 journal article on this subject became a foundational reference for virtually all subsequent researchers in this area. See Gary L Blasi, 'What Lawyers Know: Lawyering Expertise, Cognitive Science, and the Functions of Theory' (1995) 45(3) *Journal of Legal Education* 313.
[35] One prolific contributor to this field of scholarship has described this use of think-aloud verbal protocol analysis specifically as a method for the 'direct investigation of the cognitive processes of lawyers engaged in legal reasoning and problem-solving tasks.' Stratman, 'When Law Students Read Cases', above n 2, 57-58.
[36] The first of these studies was Mary Lundeberg's research on the reading skills of lawyers and law students which is discussed further below. Lundeberg, above n 2.
[37] Crombag, De Wijkerslooth and Van Tuyl Van Serooskerken, above n 3, 169. While explaining why they did not publish their earlier empirical think-aloud study findings, the authors stated that '[t]he most striking result was that what was said while thinking aloud created a rather chaotic and unsystematic impression.'

While not explicitly referencing the work of Polanyi whose treatises on tacit knowledge have been acknowledged as foundational to this field of research, [38] early non-law researchers understood both the importance and difficulties of attempting to analyse the use of tacit knowledge by legal experts. Their studies were an explicit acknowledgement of Polanyi's observation that 'we can know more than we can tell'[39] inasmuch as the attraction of think-aloud verbal protocol analysis was a direct response the so-called 'paradox of expertise' which denotes how experts commonly have difficulty explaining what it is they do and why – precisely because they are experts.[40] While traditional interviews and introspection techniques offered indirect access to cognitive processes, this alternative knowledge-elicitation and analysis methodology became popular amongst a wide range of researchers seeking to investigate directly the cognitive dimension of legal expertise and expert performance by lawyers.

The first published contribution to this field of research was a 1987 journal article by Lundeberg, a non-law academic with an interest in how law students read court cases.[41] She described in detail how the verbal protocols of experienced lawyers reading court cases differed from those of the least successful law school students.[42] Her findings informed the design of reading guidelines that were subsequently demonstrated to improve the performance of first year law students in not just reading cases, but also in terms of their overall grades. Lundeberg's research focused on the general or generic thinking skills associated with reading, not on the specialist skills required for legal practice in a specific area of law.

A number of subsequent researchers from outside the legal academy followed Lundberg's example by using think-aloud verbal protocol analysis to examine and improve the cognitive performance of law students. Their empirical research typically involved comparing and contrasting how experts and novices solved law-related

---

[38] Michael Polanyi, *The Tacit Dimension* (Doubleday, 1973); Michael Polanyi, *Personal Knowledge: Towards a Post-Critical Philosophy* (University of Chicago Press, 1974). 'Tacit knowledge is procedural knowledge that guides behavior but that is not readily available for introspection.' Robert J Sternberg, What Do We Know About Tacit Knowledge? Making the Tacit Become Explicit' in Robert J Sternberg and Joseph A Horvath (eds), *Tacit Knowledge in Professional Practice: Researcher and Practitioner Perspectives* (Lawrence Erlbaum, 1999) 231.

[39] Michael Polanyi, *The Tacit Dimension*, above n 38, 4.

[40] Crombag, De Wijkerslooth and Van Tuyl Van Serooskerken, above n 3. See also Lundeberg, above n 2, and Christensen, 'The Paradox of Legal Expertise', above n 2.

[41] Lundeberg, above n 2.

[42] This description was based on Lundeberg's doctoral dissertation: Mary A Lundeberg, *Studying Understanding in Legal Case Analysis (Reading, Self-Control Training, Metacognition)* (PhD Thesis, University of Minnesota 1985). The subject classification was 'Educational psychology.'

problems and performed generic legal tasks. Such studies have focused on legal reading skills,[43] how law students can better develop their essay-writing abilities,[44] strategies law students use to pass bar examinations following graduation,[45] students' abilities to conceptualise legal principles,[46] and how students search for, access and rely on information sources to solve legal problems.[47]

The few legal scholars who embraced these methodologies in their own empirical studies around the time that Lundeberg's study was published, such as Senger's unpublished doctoral research which analysed the extent to which law students increased their use of reasoning by analogy as they progressed through their law school studies,[48] also focused on general thinking skills.[49] The implicit assumption underlying this research was that all lawyers think the same way – or at least do so in a broad sense and insofar as matters when teaching first-year students how to pass law-school examinations.

As discussed in Part D of this chapter, some legal scholars have more recently identified thinking skills that are shared amongst legal experts, but not amongst legal novices. This is notwithstanding that all the individuals studied possessed essentially the same generic skills required to read cases and frame their arguments in an appropriate lawyerly manner. Accordingly, while generic cognitive skills have been

[43] James F Stratman, 'Contract Disclaimers in ERISA Summary Plan Documents: A Deceptive Practice?' (1988) 10 *Industrial Relations Law Journal* 350; Deegan, above n 2; Laurel Currie Oates, 'Beating the Odds: Reading Strategies of Law Students Admitted Through Alternative Admissions Programs' (1997-1998) 83 *Iowa Law Review* 139; Stratman, 'When Students Learn to Read', above n 2; Stratman, 'How Legal Analysts Negotiate Indeterminacy', above n 2; Oates, 'Leveling the Playing Field', above n 2; Christensen, 'Legal Reading and Success in Law School', above n 2; Christensen, 'The Paradox of Expertise', above n 2.

[44] Anthony Palasota, 'Expertise and the Law: Some Recent Findings from the Cognitive Sciences About Complex Human Information Processing' (1990-1991) 16 *Thurgood Marshall Law Review* 599.

[45] Sarah M Bonner and Jerome V D'Agostino, 'A Substantive Analysis of Responses to Items from the Multistate Bar Examination' (2012) 25 *Applied Measurement in Education* 1.

[46] Nievelstein, van Gog, Boshuizen and Prins, 'Expertise-Related Differences', above n 2.

[47] Helge I Stromso and Ivar Braten, 'Norwegian Law Students' Use of Multiple Sources While Reading Expository Texts' (2002) 37(2) *Reading Research Quarterly* 208; Ivar Braten and Helge I Stromso, 'A Longitudinal Think-Aloud Study of Spontaneous Strategic Processing During the Reading of Multiple Expository Texts' (2003) 16 *Reading and Writing: An Interdisciplinary Journal* 195; Nievelstein, van Gog, Boshuizen and Prins, 'Effects of Conceptual Knowledge', above n 2; Nievelstein, van Gog, van Dijk and Boshuizen, 'Instructional Support for Novice Law Students', above n 2.

[48] Charles J Senger, *Learning Legal Reasoning in Law School: The Differences Between First and Third Year Students* (PhD Thesis, Michigan State University, 1989); Charles J Senger, 'Thinking Aloud Protocols: A Diagnostic Tool for Teaching Legal Problem Solving' (1993) 10 *Thomas M Cooley Law Review* 367.

[49] Senger's doctoral research focused on the use of reasoning by analogy, which he thought senior law students may be more adept at exercising and use more frequently than junior law students when discussing legal issues, although his findings ultimately did not support this hypothesis.

observed amongst a variety of legally trained individuals as discussed above, some skills have been linked to an individual's level of legal expertise.

C *Knowledge-based Differences*

The generic conceptualisation of legal thinking skills contrasts with the ways in which cognitive studies in the field of medicine have acknowledged the specific nature of thinking skills in specialist fields of medical practice. Below is a brief review of these studies reflecting the emergence of a knowledge-based framework for comparing different levels of expertise, which has also been used in the study of lawyers. This review provides necessary background to the argument presented later that such a framework is inconsistent with the analysis of intra-specialist legal expertise, which is the focus of this thesis.

In the 1960s, Butterworth and Reppert found that experts in cardiovascular disease were consistently more skilful in identifying unhealthy heartbeats than were general practitioners, medical residents and third and fourth-year medical students.[50] Norman et al discussed similar findings in the development of dermatological expertise.[51] These latter researchers identified five levels of experts and found that specialised medical thinking skills accounted for the greater success of their higher-level experts.[52] Another team of researchers found that knowing how to read mammographic images was a skill demonstrably different from those skills developed by experienced mammographers with little image reading experience.[53]

In each of these studies 'thinking like a doctor' was of no practical or theoretical interest to researchers. Nor was it in subsequent studies concerned with specialist expertise in medicine. This contrasts with the way in which law professors have continued to characterize legal thinking skills in general terms.[54] It also contrasts with

---

[50] J Scott Butterworth and Edmund H Reppert, 'Auscultatory Acumen in the General Medical Population' (1960) 174(1) *Journal of the American Medical Association* 114.

[51] Geoffrey R Norman et al, 'The Development of Expertise in Dermatology' (1989) 125 *Archives of Dermatology* 1063.

[52] Ranked in ascending levels of expertise were (1) Second-year preclinical medical students, (2) final-year medical students, (3) Residents in family medicine, (4) General medical practitioners, and (5) Dermatologists. Ibid.

[53] Calvin F Nodine, Harold L Kundel, Claudia Mello-Thoms, Susan P Weinstein, Susan G Orel, Daniel C Sullivan and Emily F Conant, 'How Experience and Training Influence Mammography Expertise' (1999) 6(10) *Academic Radiology* 575.

[54] Examples of textbooks in this category, include: Sanson, Anthony and Worswick, above n 9; Patrick Keyzer, *Legal Problem Solving – A Guide for Law Students* (LexisNexis, 2002); Romantz and Vinson,

the critical-thinking literature produced by legal scholars[55] and, as has been noted, with surveys of legal practitioners which perpetuate the view that 'to think like a lawyer' means having been taught generic cognitive skills at law school on the implicit understanding that these skills will be foundational to future success in legal practice.

In the early 1990's, two leading cognitive psychologists working in the field of medical expertise moved beyond the expert-novice dichotomy of cognitive analysis to develop a knowledge-based hierarchy of expertise.[56] Patel and Groen described a conceptual framework within which it was 'possible to consider much finer gradations of expertise … between specific expertise (e.g., cardiology) and generic expertise (e.g., medicine)' on the basis that 'an individual may possess both, or only generic expertise.'[57] The authors further explained by way of an example from an unrelated field that 'in computer programming, systems programmers are a different breed from applications programmers,' even if they might both be regarded somewhat simplistically as just 'computer programming experts' by general observers.[58] They also noted that expertise based on specialization 'is apparent in every academic discipline,' and that 'in chess, masters specialize in different opening variations.'[59]

Patel and Groen's knowledge-based hierarchy of expertise is reproduced in Table 2.1. In this table, there is a conceptualized progression from layperson to expert. But it is not a progression within the same specialist knowledge domain. Rather, it is the

---

above n 9; John Farrar, *Legal Reasoning* (Thomson Reuters, 2010); Larry Alexander, *Demystifying Legal Reasoning* (Cambridge University Press, 2008); Bartosz Brozek and Jerzy Stelmach, *Methods of Legal Reasoning* (Law and Philosophy Library, 2006); Richard Posner, *How Judges Think* (Harvard University Press, 2008); Kenneth J Vandevelde, *Thinking Like a Lawyer: An Introduction to Legal Reasoning* (Westview Press, 1996).

[55] Legal thinking in a generalised or generic sense has been variously equated with critical thinking, logical thinking, non-assumptive thinking, legal problem solving, legal analysis, legal reasoning, how judges think, and creative thinking. A list of all alternative and equivalent concepts discussed in the literature on legal thinking skills is too long to document comprehensively here. For an overview of this literature see Sanson, above n 21. See also James, above n 9, in which the author notes the Australian Learning and Teaching Council's Threshold Learning Outcome (TLO3) for the Bachelor of Laws degree titled 'Thinking Skills' and how this TLO3 was formulated to align with relevant studies, reports and recommendations published in a range of other jurisdictions, including Britain, the United States, Canada and Scotland. These related documents (which are cited by James) include extensive discussions on what it may mean to think like a lawyer from educational, institutional and professional perspectives.

[56] Vimla L Patel and Guy J Groen, 'The General and Specific Nature of Medical Expertise: A Critical Look' in K Anders Ericsson and Jacqui Smith (eds) *Toward a General Theory of Expertise: Prospects and Limits* (Cambridge University Press, 1991) 93.

[57] Ibid 96.

[58] Ibid.

[59] Ibid. For a recent study of the differences between the approaches of expert chess players who specialise in different opening strategies, see Merim Bilalic, Peter McLeod and Fernand Gobet, 'Specialization Effect and Its Influence on Memory and Problem Solving in Expert Chess Players' (2009) 33 *Cognitive Science* 1117.

possession of specialized knowledge that distinguishes experts from less able problem-solvers. This results in a hierarchy which is purely knowledge-based and conforms to a binary 'jump' analogy rather than to a graduated progression in proficiency. This is most clearly seen in the distinction between subexperts and experts, where the only difference is that the latter possesses the relevant specialist knowledge, but the former does not.

TABLE 2.1 – Patel and Groen's Knowledge-Based Hierarchy of Expertise[60]

Terminology	Definition
Layperson	An individual who has only common-sense or everyday knowledge of the domain.
Beginner	An individual who has the pre-requisite knowledge assumed by the domain.
Novice	A layperson or a beginner.
Intermediate	By default, we define this as anyone who is above the beginner level but below the subexpert level.
Subexpert	An individual with generic knowledge, but inadequate specialized knowledge, of the domain.
Expert	An individual with specialized knowledge of the domain.

A small group of legal scholars have undertaken cognitive research consistent with – and in some instances explicitly based on – this view of how experts think differently from non-experts. Mitchell was the earliest to record his findings in this area,[61] although his research was not presented as a scientific study.[62] Having been influenced by research from the social sciences in which think-aloud verbal protocol analysis had revealed how experts in Soviet Union Government policies solved a hypothetical Soviet grain supply problem more effectively than experts from other fields, such as chemistry, [63] he observed how his law-school colleagues with specialist expertise in criminal law were far more effective in solving a criminal law problem than law professors who specialised in other areas of law.

[60] Patel and Groen, above n 56, 96.
[61] John B Mitchell, 'Current Theories on Expert and Novice Thinking: A Full Faculty Considers the Implications for Legal Education' (1989) 39 *Journal of Legal Education* 275.
[62] Mitchell did not use think-aloud verbal protocol analysis but rather an informal form of introspective interviews. Regarding the intellectual progress made during the exercise he described in his article, the author noted, 'The results are interesting but hardly lay claim to being scientific.' Ibid 281-282.
[63] James F Voss, Telly R Greene, Timothy R Post and Barbara C Penner, 'Problem-Solving Skill in the Social Sciences,' in Gordon H Bower (ed) *The Psychology of Learning and Motivation* (New York, 1983) 165.

Mitchell noted that although his 'criminal law or "expert" group approached such a problem in a number of different ways, each approach had common characteristics.'[64] Because each group member had appropriate knowledge-based schema and methods for dealing with problems of this kind, they were able to apply themselves to the task quicker than their colleagues. Mitchell stated, 'In short, the experts "saw" – or could construct – a coherent whole that was triggered by and transcended the facts, into which the facts could be fit[ted] to arrive quickly at an overall solution.'[65] He also recorded how his group of criminal law professors 'worked with apparent ease and enjoyment.'[66]

These findings contrasted with the performance of Mitchell's non-experts. These individuals were frustrated insofar as they knew what was expected of them and what the end result of their deliberations should have been. But they lacked the key ingredients to complete the task because they did not have adequate doctrinal and policy knowledge. They were therefore unable to develop a reasoned and satisfactory answer, or at least not do so as easily as their criminal law colleagues could. Mitchell concluded that these non-experts, while not lacking in analytic ability, 'simply did not know enough.'[67] This is notwithstanding that these non-experts actually were experts, but in other areas of law.

Marchant and his fellow business-school colleagues sought to compare the analogical reasoning abilities of tax accountants with those of undergraduate tax students in an experiment involving a tax law problem.[68] While the authors had inadvertently confounded their original experiment by introducing a hypothetical legal rule that conflicted with the tax deduction rules with which their experts were familiar,[69] the fact that they had recognised taxation law as involving specific cognitive skills was novel in the context of then existing literature on legal thinking skills.

---

[64] Mitchell, above n 61, 282.
[65] Ibid.
[66] Ibid.
[67] Ibid 283.
[68] Garry Marchant, John Robinson, Urton Anderson and Michael Schadewald, 'Analogical Transfer and Expertise in Legal Reasoning' (1991) 48 *Organizational Behaviour and Human Decision Processes* 272.
[69] These researchers were therefore unable to show that the more expert advisors in this study had better analogical transfer skills than less expert students. In fact, the researchers found that the novices performed better because they had not internalized the tax deduction rule that the experts could not get out of their minds as required to succeed at the representative task.

However, while Marchant et al had created a protocol coding system to trace the reasoning processes of their experiment subjects, they did not apply this system to the transcripts of think-aloud concurrent verbalisations. Instead they relied on written answers to a test problem, which did not permit the kinds of insights into actual thinking skills that are possible with think-aloud verbal protocol analysis. Nevertheless, their conception of an intra-specialist test of reasoning strategies in a legal setting was novel, even if their comparison of novices and domain experts was not.

D *Specialist Legal Expertise*

The earliest published statement of interest by a legal scholar in the use of think-aloud verbal protocol analysis to study the thinking skills of experts in specialist fields of legal practice, appeared in a 1995 journal article by law professor Gary Blasi.[70] He was the first legal academic to state explicitly that cognitive science provided the 'theoretical framework and empirical methods that make it possible to study directly a topic about which law professors have long only asserted knowledge: how lawyers think.'[71] With references to the work of leading cognitive scientists such as Ericsson and Simon, authors of the seminal text *Protocol Analysis: Verbal Reports as Data* ('Protocol Analysis'),[72] Blasi claimed that law professors could at last investigate, 'topics until now considered either simply unknowable or suitable only for speculation or mere assertion: judgment, wisdom, expertise and the relation of theory to lawyering practice.'[73]

Like Mitchell, Blasi's optimism was primarily based on the results of the Soviet Union grain supply study published in 1983 as well as on the work of cognitive psychologists in the medical field. He also referenced other studies including one published in 1993 by Schraagen, whose expertise as a psychologist with an interest in experimental design resulted in a textbook study of specialist expertise in gustatory research.[74] Blasi did

[70] Blasi, above n 34.
[71] Ibid 354.
[72] Anders K Ericsson and Herbert Simon, *Protocol Analysis: Verbal Reports as Data* (The MIT Press, 1984/93).
[73] Blasi, above n 34, 317.
[74] Jan Maarten Schraagen's experiment compared the cognitive differences between experts in the design of psychology experiments used in gustatory research, and experts whose expertise was in designing experiments in sensory psychology, but who had not designed experiments for gustatory research. Jan Maarten Schraagen, 'How Experts Solve a Novel Problem in Experimental Design' (1993) 17(2) *Cognitive Science* 285.

not, however, report on any of his own studies involving think-aloud verbal protocol analysis of legal expertise, although he did indicate that he was undertaking such research.[75] Instead, he offered an extended discussion on the need for such empirical research amongst his law school colleagues, noting in the process that he was making a contribution 'to an academic discourse that does not yet exist.'[76]

Blasi claimed that, 'despite an enormous amount of theoretical and empirical work on expertise in problem-solving in dozens of different domains, none of that work to date has examined expert problem-solving by lawyers.'[77] He then offered a tentative definition of legal expertise as 'extraordinary competence in the instrumental solving of the problems of clients.'[78] Subsequent researchers have cited Blasi as providing '[a]n excellent discussion of the implications of cognitive psychology research for legal education and legal problem-solving'[79] and that his 'rich and insightful article [set] clinical scholarship in the cognitive science framework.'[80]

Two legal scholars directly influenced by Blasi's article were Colon-Navarro and Weinstein. These scholars published empirical studies of a kind broadly comparable to the approach envisaged by the present research project.[81] Krieger, another legal academic who explicitly embraced Blasi's perspective on the value of cognitive research into how lawyers think and who was heavily influenced by the work of researchers such as Patel and Groen in the medical field, also undertook think-aloud studies but focused on law students rather than specialist legal practitioners.[82]

---

[75] Blasi records in footnote 127 of his article his own use of think-aloud verbal protocol analysis in an unpublished (and possibly incomplete) study of the cognitive differences between a novice lawyer and a lawyer with two decades of experience in an unspecified area of law. Blasi, above n 34.
[76] Ibid 320.
[77] Ibid 318. Blasi did not seem to view the work of non-lawyers, such as Mary Lundeberg and the other authors noted above (above n 2), as involving 'expert problem-solving by lawyers.' This may have been because he did not consider general or generic legal skills to be the same as expert legal skills in a more specialist sense. Or he may have just limited his review of relevant literature to that published by legal academics.
[78] Ibid 320.
[79] Stefan H Krieger, 'Domain Knowledge and the Teaching of Creative Legal Problem Solving' (2004-2005) 11 *Clinical Law Review* 149, 164.
[80] Weinstein, above n 2, 7.
[81] The broad similarities between these two studies and the present study mainly relate to their shared acknowledgement of specialist legal expertise as a discrete area of cognitive research and their utlisation of think-aloud verbal protocol analysis. Their use of novice-expert tests, where specialist domain knowledge is the difference at issue, distinguishes them from the intra-specialist tests developed for this thesis.
[82] Stefan H Krieger, 'The Development of Legal Reasoning Skills in Law Students: An Empirical Study' (2006) 56(3) *Journal of Legal Education* 332. Krieger's study compared the reasoning strategies of

Colon-Navarro described research he had undertaken involving immigration law experts solving an immigration law problem.[83] His study focused on differences between how his selected experts interviewed a client with an immigration law issue compared with how two sets of novices approached the same task.[84] One of these sets of novices consisted of law students who had both theoretical and clinical training in immigration law. The other was a lone law student with only theoretical training in this area of specialist legal knowledge.[85]

Modelling his methods after those 'used by cognitive psychologists studying the development of expertise in the medical field as well as in chess and physics,' Colon-Navarro devised a legal problem using an actor to play the role of a man from El Salvador seeking to live in the United States.[86] His expert subjects were better at sorting between the relevant and irrelevant facts embedded in the problem, they also developed workable plans by the ends of their interviews (as well as alternative plans), they acted more confidently and reassuringly towards their client, and they prioritised the available remedies in a similar way to each other. The novices in the study, on the other hand, had a different ordering approach (though the study's experienced novices were more like the experts in this regard) [87] and were much slower at 'diagnosing' the problem and coming up with solutions.[88]

Colon-Navarro's concluding remarks echoed the informal findings of Mitchell. He noted how his immigration law experts used their better organized substantive and procedural knowledge to identify a range of possible solutions, as well as how they

---

incoming law-students, second-year law students and final-year law students, who were all required to consider a consumer fraud problem relating to the sale of a used car.

[83] Colon-Navarro, above n 2.

[84] The task was to assess the effectiveness of clinical legal education in enabling a novice lawyer to begin 'developing the "mental file" that the experienced attorney has available to himself or herself.' It was also intended that the results of the study be used 'to compare "legal expertise" to expertise in other fields (as described in the cognitive psychology literature).' Ibid 120.

[85] Ibid 121-122.

[86] A detailed biography and situation description of this hypothetical client as given to the actor was set out over almost four pages of Colon-Navarro's article, starting at page 123. An actor was used for ethical reasons and to ensure that the author could maintain strict control over the content of the issues, including the placement of 'red herrings' within the factual matrix of the case. Ibid 123.

[87] Ibid 128-132.

[88] 'A very striking difference between the experts and novices was the speed at which the experts reached the hierarchical ordering of the obtainable remedies and were ready to counsel the client. Within five minutes of the interview's start, the experts were clearly focusing on the client's relationship with Maria ("establishing a bona fide," in one expert's words) and bringing up the question of marriage to Maria. While this may seem uncomfortably swift to some, the fact that it was displayed by all the experts points to it being a significant finding, the meaning of which is certainly worthy of further study.' Ibid 129.

'looked at the facts (analogous to the 'x-rays' in the study of radiologists) and quickly extracted only relevant aspects, leading to the correct diagnosis.'[89]

Weinstein began his 1998 journal article by claiming that 'the doctrine-centric conception of what it means to "think like a lawyer"' had blinded legal academics to the 'important differences in the ways lawyers approach different kinds of problems' and that it is a mistake to 'confuse appellate lawyering for all of lawyering.'[90] He then explained how his own studies had provided 'a model of how legal and factual command over a specific area of law allows lawyers to quickly, and largely unreflectively, analyze both facts and law together.'[91] He further claimed that 'the degree of specialized knowledge required may surprise those who view "good lawyering" as a generic ability, applicable to any legal question.'[92]

Weinstein, like Colon-Navarro, drew directly on the research of cognitive psychologists working in the medical field. In relation to his contention that area-specific knowledge in law is as critical to problem solving expertise in law as it is in other areas,[93] he confirmed his view that: 'The relevant area of knowledge is usually narrow, often a professional subspecialty. For doctors it would be a particular area of medicine such as cardiac surgery or radiology, not medicine generally.'[94]

Weinstein contended that his study constituted 'the first application of the human problem solving model to representational lawyering.'[95] He also claimed that it was conceived as a direct response to Blasi's challenge to apply 'the human problem solving model to lawyering.'[96]

---

[89] Ibid132.
[90] Weinstein, above n 2, 3.
[91] Ibid.
[92] Ibid 3, footnote 4.
[93] Ibid 6. 'Studies using the model [of human problem solving] in other disciplines show that experienced thinkers knowledgeable in a particular area use different cognitive operations to solve problems in their areas of expertise than do those lacking specific experience. The expert's cognitive techniques are characteristically fast and unconscious. This study shows that the same is true in law.' Ibid 5-6 (references omitted).
[94] Ibid 5, footnote 15.
[95] Ibid 5. The reference to 'representational lawyering' distinguished Weinstein's study from the study of magistrates engaging in sentencing drink-driving, shoplifting and drug-related offenders, concerning which verbal protocols had been previously analysed. See Jeanette A Lawrence, 'Expertise on the Bench: Modeling Magistrate's Judicial Decision-Making' in Michelene T H Chi et al (eds) *The Nature of Expertise* (Lawrence Erlbaum Associates, 1998), which Weinstein claimed was the first application of the human problem solving model to law.
[96] Weinstein, above n 2, 9.

Weinstein's think-aloud verbal protocol analysis centred on the case assessment of a fictitious SSD claimant called Mr Sims.[97] 'SSD' refers to Social Security Disability, and the experimental task required participants to help Mr Sims solve the problem of his initial claim having been rejected by the Social Security Agency. Mr Sims' case, participants were told, had been prepared by legal services officers who had arranged for him to appear before an Administrative Law Judge. According to comments made by one of the expert lawyers in the experiment, this was typical of case files he handled at his office.[98]

There were 10 participants in Weinstein's study. Three were described as 'outstanding practitioners' by the author, although their experience and credentials were not detailed.[99] They were designated experts by virtue of their experience in dealing with SSD claims. There was one subexpert, whom the author described as a legal academic 'experienced in appellate and trial level criminal law,' but who had little experience with SSD claims.[100] The remaining six participants were law students who had some familiarity with SSD cases, including in some instances representing 'an SSD claimant in a live client clinic.'[101]

The lone legal academic's status as a subexpert was based on his being an expert in a related field of law, but lacking 'specific knowledge in the domain under examination.'[102] This was an explicit reference to the work of Patel and Groen,[103] developers of the knowledge-based hierarchy of expertise in Table 2.1 above, in their research on the problem-solving abilities of specialist medical professionals.[104] As noted, these researchers had defined a subexpert as: 'An individual with generic knowledge, but inadequate specialized knowledge, of the domain.'[105]

In summing up his assessment of the subexpert's performance, based on an analysis of this participant's think-aloud verbal protocols, Weinstein observed that while the appellate attorney was very capable at doctrinal analysis, he lacked 'both substantive

[97] A detailed summary of the factual and legal background to the case is provided in ibid 20-23.
[98] Ibid 19, footnote 89.
[99] Ibid 18.
[100] Ibid.
[101] Ibid 19.
[102] Ibid 18, footnote 88.
[103] Ibid 24, footnote 119.
[104] Patel and Groen, above n 56.
[105] Ibid 96.

knowledge and domain-specific models for dealing with the much more ill-structured problem presented by initial case planning [of SSD claims].'[106] Nevertheless, this experienced lawyer was able to develop a compelling solution to the set task. The only problem was, according to Weinstein, his conclusion 'just happened to be wrong.'[107]

E *Comparing Comparative Studies*

A critical dimension to the above literature is the comparative nature of previous expertise studies. Common to virtually all the studies noted has been a focus on how experts think differently from novices. Studies of this kind have been by far the most common amongst researchers in this area, ostensibly reflecting a motivation to improve the performances of novices (particularly undergraduate law students) by better understanding how their cognitive skills differ from (and may be improved through pedagogical intervention to be more like) those of presumed experts.

Expert-subexpert studies have been less common, at least in relation to legal thinking skills. Whereas expert-novice research has mostly compared students with professionals, expert-subexpert tests have tended to compare professionals at the same general level of expertise and experience – but with differing amounts of the specific knowledge critical to performing the required problem-solving tasks.

What have not been studied in the legal context are cognitive differences associated with different levels of expertise within the *same* area of domain specialisation. The conceptual distinction between this approach and the more commonly followed approaches described above can be explained with reference to the following comparison of two alternative schemes for categorizing specialist expertise. The first scheme is the knowledge-based hierarchy used by Patel and Groen described in Table 2.1. The second is the traditional, progressive categorization of expertise described by Hoffman.[108] Table 2.2 below sets out both the Patel and Groen's framework and what will be termed Hoffman's Scheme[109] in a side-by-side comparison.

---

[106] Weinstein, above n 2, 52.

[107] Ibid. Weinstein went on to note: 'The subexpert, a very talented lawyer, recognized his limits and was reluctant to analyse the case because he lacked domain-specific knowledge. Not all lawyers, however, recognize their limits. The plausibility of the subexpert's answer suggests some of the dangers of putting too much faith in the doctrine-centric view of thinking like a lawyer.' Ibid footnote 221.

[108] Robert R Hoffman, 'How Can Expertise be Defined? Implications of Research from Cognitive Psychology' in Robin Williams, Wendy Faulkner and James Fleck (Eds), *Exploring Expertise: Issues and Perspectives* (Macmillan, 1998) 84. Hoffman's Scheme has been more recently presented and relied

TABLE 2.2 – Comparison of Alternative Expertise Categorisation Schemes

Patel & Groen's Knowledge-based Hierarchy [110]	Hoffman's Scheme – A Traditional Progressive Categorisation of Expertise
Layperson: *An individual who has only commonsense or everyday knowledge of the domain.* **Beginner:** *An individual who has the pre-requisite knowledge assumed by the domain.* **Novice:** *A layperson or a beginner.* **Intermediate:** *By default … anyone who is above the beginner level but below the subexpert level.* **Subexpert:** *An individual with generic knowledge, but inadequate specialized knowledge, of the domain.* **Expert:** *An individual with specialized knowledge of the domain.*	**Naivette:** *One who is totally ignorant of a domain.* **Novice:** *Literally, someone who is new – a probationary member. There has been some ('minimal') exposure to the domain.* **Initiate:** *Literally, someone who has been through an initiation ceremony – a novice who has begun introductory instruction.* **Apprentice:** *Literally, one who is learning – a student undergoing a programme of instruction beyond the introductory level. Traditionally, the apprentice is immersed in the domain by living with and assisting someone at a higher level. The length of an apprenticeship depends on the domain, ranging from about one to 12 years in the craft guilds.* **Journeyman:** *Literally, a person who can perform a day's labour unsupervised, although working under orders. An experienced and reliable worker, or one who has achieved a level of competence. It is possible to remain at this level for life.* **Expert:** *The distinguished or brilliant journeyman, highly regarded by peers, whose judgments are uncommonly accurate and reliable, whose performance shows consummate skill and economy of effort, and who can deal effectively with certain types of rare or 'tough' cases. Also, an expert is one who has special skills or knowledge derived from extensive experience with sub-domains.* **Master:** *Traditionally, a master is any journeyman or expert who is also qualified to teach those at a lower level. Traditionally, a master is one of an elite group of experts whose judgments set the regulations, standards, or ideals. Also, a master can be that expert who is regarded by other experts as being 'the' expert, or the 'real' expert, especially with regard to sub-domain knowledge.*

As noted, Patel and Groen's framework was used by Weinstein to assess cognitive differences between the experts and subexpert in his experiment involving SSD claims. The fact that the subexpert in Weinstein's study differed from the designated experts only inasmuch as he lacked direct specialist knowledge of immigration law cases of a particular kind, fits squarely within the Patel and Groen knowledge-based hierarchy. It

---

on in Robert R Hoffman, Paul Ward, Paul J Feltovich, Lia DiBello, Stephen M Fiore and Dee H Andrew, *Accelerated Expertise: Training for High Proficiency in a Complex World* (Psychology Press, 2014).

[109] Hoffman has commented on this traditional, proficiency-based approach to ranking expertise, as follows: 'This scheme has been faulted for its reliance on an outdated, male-oriented perspective, and yet alternative terminological schemes have not been forthcoming. If one acknowledges that expertise develops, and that qualitative changes occur over the developmental period, then one must make some attempt at a stage-like categorization, if only to motivate research. I would welcome alternative conceptual schemes and am myself not religiously attached to the diminutive "naivette" neologism.' Hoffman, 'How Can Expertise be Defined?' above n 108, 95 endnote 1. More recently, Michelene Chi has confirmed the utility of this scheme in her authoritative review of different approaches to expertise studies. See M T H Chi, 'Two Approaches to the Study of Experts' Characteristics,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 21.

[110] Patel and Groen, above n 56, 96.

does not, however, comport with Hoffman's traditional progressive categorization scheme, which is described by Chi as a 'proficiency scale.'[111] This is because the only difference between Weinstein's experts and subexpert was that the former had the required specialist domain knowledge while the latter did not.

It is not feasible under Patel and Groen's framework to compare an expert and a lesser expert who both possess essentially the same domain-specific knowledge. This is because knowledge is the differentiating variable that separates the two categories of experts. It may be contended that an expert has *more* specific knowledge in the given domain. But this is inconsistent with a scheme where one must either have knowledge or not have knowledge rather than varying amounts of it – at least not in relation to the specialist knowledge of the relevant expertise domain. This is the binary essence of Patel and Groen's conceptual framework.

Hoffman's Scheme, on the other hand, presumes a progression of knowledge and capabilities associated with different amounts of instruction and domain-specific experience. It describes relative levels of expertise or proficiency within a single knowledge domain. In doing so, it fulfils its traditional function as a means of describing the progression of increasing knowledge and skills as one moves, over many years, from the status of novice to a master in a specific field.

Consider for instance a team or department of specialist competition law lawyers as exists in most large commercial law firms, with, say, three partners, two special counsel and five senior associates,[112] or alternatively seven partners and 30 other lawyers.[113] Each lawyer in these teams specialises in competition law exclusively or to a significant degree, yet they are recognised as having different levels of skill or problem-solving ability. This is the case even though, at least in a technical or an academic sense, they might all possess essentially the same amount of technical knowledge about this specialist area of law. Such a team will often include esteemed

[111] Chi, above n 109, 22.
[112] These are the data provided for DLA Piper in Australia (http://www.dlapiper.com/en/australia), whose competition law practice is positioned in Band 4 of *Chambers and Partners'* ranking of Australia's leading competition law practices. http://www.chambersandpartners.com/23/26/Editorial/8/1#22401592_editorial accessed 14 May 2014.
[113] These are the data provided for Allens in Australia (www.allens.com.au), whose competition law practice is positioned in Band 2 of *Chambers and Partners'* ranking of Australia's leading competition law practices. http://www.chambersandpartners.com/23/26/Editorial/8/1#22401592_editorial access 14 May 2014.

experts working alongside more junior solicitors, as well as senior lawyers on the cusp of making partnership.[114]

From a dollar-per-hour fee perspective (where the most senior lawyers in such a specialist legal team may charge hourly fees more than double those of their junior colleagues[115]) and in terms of the allocation of formal responsibilities within the group (partners having sign-off authority,[116] while non-partners typically not), it is much easier to explain these differences under Hoffman's Scheme than by attempting to rely on Patel and Groen's framework. In practical terms, one would ostensibly be hard-pressed to establish the necessary differences in knowledge to justify calling any of these lawyers 'subexperts' or, to use Patel and Groen's description of the term, individuals 'with generic knowledge, but inadequate specialized knowledge, of the domain.'[117]

F *Intra-Specialist Expertise Comparisons*

The preceding discussion considered the few empirical studies that have investigated the cognitive differences between how expert and less expert practitioners in the same specialist field of law solve domain-specific problems. The farthest any researcher has come in answering this thesis' research question is Weinstein. However, his conception of subexpertise was not of a progressive or staged proficiency, but a knowledge-based one within Patel and Groen's framework. The alternative framework of Hoffman's Scheme, which as noted in the previous section traces the advancement of apprentices to journeymen and experts to masters, envisages an intra-domain progression where comparisons are made not between individuals with and without

[114] According to *Chambers and Partners'* ranking methodology, three of Allens competition law specialists are ranked in Band 2, one is in Band 3 and two are in Band 4. Based on previously noted statistics about the make-up of the Allens' competition law department (see above n 113), this leaves one partner-level lawyer and another 30 lawyers neither ranked nor cited in *Chambers and Partners'* directory, all of whom may be assumed to possess different levels of specialist expertise in competition law.

[115] A recent study of lawyers' hourly rates in London commission by the Legal Services Board noted: '[H]ourly rates vary substantially by qualifications and the length of experience with rates for partners often being multiples of the rates charged by more junior lawyers.' Charles River Associates, *Benchmarking the Supply of Legal Services by City Law Firms*, report prepared for Legal Services Board, August 2011, 66. Figure 13 of that report indicates that partners in some Magic Circle and US law firms can charge up to 50% more than lawyers with 5 years post-qualification experience.

[116] Sign-off authority as a concept is described in greater detail the next chapter. Briefly, it connotes an individual having the authority to sign official documents and legally bind their organization. An example is a law firm partner signing a letter of legal advice on behalf of their firm.

[117] This quotation is part of the definitions of Patel and Groen's subexpert category as set out in Table 2.2 above.

domain knowledge, but between individuals with different levels of domain ability which may or may not be attributable to differences in domain-specific knowledge.

However, Hoffman's use of traditional definitions of expertise based on Middle Ages' craft guilds in Europe[118] may be criticised for its apparent lack of scientific rigour and motivate a search for more modern conceptual frameworks. It is therefore useful to examine the practical utility of this terminology more closely, as well as to consider alternatives and their suitability for use in this thesis.

Considering journeymen first, the essential character of such an individual is that he or she must be capable of working unsupervised in a specialist type of work. This requires that they possess both considerable technical competency and an appropriate level of trustworthiness. Moreover, the fact that one may remain a journeyman for life as Hoffman notes, suggests that (i) it is a vocation which has traditionally been sufficiently compensated so as to provide a living wage (in other words, it is not a part-time or hobby activity demanding only a casual interest), and (ii) the status of a journeyman is not necessarily associated with youth. Indeed, a degree of maturity would seem essential to fulfil the role as defined.

Critically, however, a journeyman is also an individual who works under orders. In other words, he or she requires direction of some kind, presumably a higher-level of direction than that associated with mere supervision. This also suggests that a journeyman would not be held ultimately responsible for whether their finished product is necessary or appropriate – just that they complete their work as instructed. In this sense, they could be viewed as part of middle-management in a contemporary setting.

Experts, on the other hand, have their status determined by being 'highly regarded by their peers,' having 'uncommonly accurate and reliable' judgment, demonstrating 'consummate skill and economy of effort,' and being able to 'deal effectively with certain types of rare or "tough" cases.'[119] They may also demonstrate 'special skills or knowledge derived from extensive experience with sub-domains.'[120] Experts, it may therefore be assumed, would be capable of giving orders to journeymen and to lower level workers. In this capacity they would not only make decisions as to what needs to

[118] Hoffman, 'How Can Expertise Be Defined?' above n 108.
[119] These quotations are part of the definitions from Hoffman's Scheme as set out in Table 2.2 above.
[120] Ibid.

be done and how, but would take responsibility for those decisions in ways that a journeyman would not.

The issue of rare or tough cases is of particular interest inasmuch as such cases may occur only infrequently, making it possible for a legal specialist to work for several years in their area of law without ever encountering such cases. As noted by Hoffman in his account of operators of large industrial looms in the mid-1900s,[121] an expert operator was generally required to have at least seven years' experience in this field since some atypical events – notably certain loom-breakdown scenarios – occurred, on average, only once every seven years.[122] This led trainers to develop a loom that was designed to breakdown more frequently, thereby speeding up the development of operator expertise.[123]

An expert is therefore likely to have more experience, as measured in time, as this is necessary to give them exposure to unusual tests of their cognitive skills – unless of course special training is provided as in the case of the loom operators. It is also noted how Norman et al's research in the medical field found that specialists in dermatology were better than less experienced colleagues – who were still thoroughly competent dermatologists in their own rights – in relation to atypical cases.[124] In that study, this was found to be the only way of distinguishing between the participating experts and lesser experts in terms of their cognitive performances.

Regarding the definition of a master under the above framework, attention is drawn to whether a legal specialist performs a formal or informal teaching role (perhaps at the post-graduate level where specialised, craft-like training is most likely in a modern university context), whether they are viewed as part of an elite group of standard setters, and how other experts in the field assess their abilities. Where the activities of an individual satisfy the requirements of being an expert and these additional indicators are evident, a relatively confident assessment of mastery under Hoffman's Scheme would seem feasible.

[121] Hoffman, 'How Can Expertise be Defined?' above n 108, 93.
[122] Ibid 94.
[123] J J Jenkins, 'Educating for Applications' in R R Hoffman and D S Palermo (eds), *Cognition and the Symbolic Processes: Applied and Ecological Perspectives* (Erlbaum, 1983) 335.
[124] Norman et al, above n 51.

The remaining categories of naivette, novice, initiate and apprentice are also capable of objective assessment in a general sense. By defining the relevant knowledge domain as a specialist field of law, for instance competition law as above, whether someone is totally ignorant of the laws in this area, has had minimal exposure to it, has been given introductory instruction, or is 'undergoing a program of instruction beyond the introductory level' and may be considered 'immersed in the domain' insofar as they are working in a team where they provide assistance to higher level specialists, are all questions that could be answered objectively with only limited investigation. The appeal of such descriptors is clear in an empirical setting where distinctions between higher levels of specialist expertise are important.

Hoffman's endorsement of this analytical framework in the absence of better alternatives[125] is further confirmation of its utility for this thesis. Using Taleb's arguments regarding time and fragility, the fact that a framework of this kind has been in continuous use for several centuries (largely in practical and therefore empirically rigorous contexts) is itself testament both to its theoretical robustness and its relevance for centuries to come.[126] It is also noteworthy how the legal profession today maintains remnants of the traditional craft guild framework insofar as clerkship and pupillage are still undertaken under the instruction and supervision of a 'master' lawyer, and as has been discussed, specialist teams of lawyers in law firms typically include members who may be placed conceptually with little difficulty in their appropriate categories within Hoffman's Scheme.

In addition to Hoffman's Scheme and Patel and Groen's knowledge-based distinction between experts and subexperts (which Weinstein adapted for his research but which is ill-suited to the present study), Chay has identified three further conceptual frameworks for assessing and categorising different levels of expertise on a progressive basis.[127] However, none of these appears to add much to Hoffman's Scheme – and could involve

[125] Hoffman, 'How Can Expertise be Defined?' above n 108, 95 endnote 1.
[126] '[T]he old is superior to the new … No matter how something looks to your intellectual machinery, or how well or poorly it narrates, time will know more about its fragilities and break it when necessary …' and 'If a book has been in print for forty years, I can expect it to be in print for another forty … Every year that passes without extinction doubles the additional life expectancy. This is an indicator of some robustness. The robustness of an item is proportional to its life.' Nassim N Taleb, *Antifragile: How to Live in a World We Don't Understand* (Allen Lane, 2012) 309 and 318.
[127] Allan James Chay, *Lawyer Problem Solving: An Investigation of the Knowledge Used in Solving Practical Legal Problems* (PhD Thesis, Griffith University, 2006) 80-81.

serious feasibility issues in the present context while at the same time being viewed as less well established frameworks of progressive expertise categorisation.

Dreyfus and Dreyfus,[128] who developed their framework while researching activities as varied as learning languages, playing chess and piloting aircraft, identified five stages or phases of expertise development: (i) Novice (able to correctly pair problems and problem-solving rules); (ii) Advanced beginner (able to successfully undertake the same pairing exercise as at the novice level, but in a wider range of contexts); (iii) Competence (able to consider an even wider range of problems in real-world settings, but lacking 'a sense of what is important'[129]); (iv) Proficiency (able to reflect on alternative courses of action while starting to act intuitively based on previous experiences); and, (v) Expertise (being able to rely more on intuition without needing to make conscious decisions, but simply doing 'what normally works'[130]).

The other two frameworks for categorising expertise development identified by Chay were those of Schmidt, Norman and Boshuizen, whose concepts were derived from their work with medical practitioners,[131] and Scandura, who was concerned primarily with the efficacy of general school education.[132] The former researchers' framework took the form of a four-stage model in which experts were distinguishable from lower-level practitioners by their development of illness scripts. These scripts were defined as 'memories of previous patients [that] are retained in memory as individual entities and are not merged into some prototypical form.'[133] The latter researcher favoured a three-stage analytical framework consisting of naïve, neophyte and master, with mastery being reflected in greater efficiency and the use of more complex cognitive structures.

Chay dismissed each of these frameworks for his study because they required an assessment of cognitive development, which depended on investigations of individual's cognitive abilities in a manner more extensive than he was prepared to undertake.[134]

---

[128] HL Dreyfus and SE Dreyfus, *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer* (The Free Press, 1986).
[129] Ibid 23.
[130] Ibid 31.
[131] H G Schmidt, G R Norman and H P A Boshuizen, 'A Cognitive Perspective on Medical Expertise: Theory and Implications' (1990) 65(10) *Academic Medicine* 611.
[132] JM Scandura, 'Problem Solving in Schools and Beyond: Transitions from the Naïve to the Neophyte to the Master' (1981) 16(3) *Educational Psychologist* 139.
[133] Schmidt, Norman and Boshuizen, above n 131, 617.
[134] Chay, above n 127, 81.

In any event, he concluded that his study was not dependent on establishing stages of development, but rather on what he termed 'transformations of knowledge.'[135] The objective criteria of Hoffman's Scheme may have been a more practical approach for him to have adapted, to the extent that a degree of staged categorisation was required to address his stated research questions. In this respect, Chay's ultimate reliance on years of experience and peer recognition to identify his experts was broadly consistent with – even if not based on – the traditional framework described by Hoffman.

G *Conclusion*

The above discussion confirms that no previous empirical studies provide a complete response to this study's research question which is concerned with how legal specialists with different levels of expertise (but the same technical legal knowledge) think differently from one another. This is notwithstanding the extensive research and numerous studies that have been undertaken by scholars both within and outside the legal academy which constitute both a rigorous and generous base on which to construct an appropriate methodology for the present research. This includes: the think-aloud problem solving and verbal protocol analysis methodologies that have been developed and refined by cognitive psychologists, and which have been used in the study of legal thinking skills; and Hoffman's Scheme, which relies on a well-established conceptualisation of progressive expertise and is analytically aligned with the objectives of this thesis.

The research undertaken by Colon-Navarro and Weinstein offers useful technical guidance, and the expertise traits observed in the expert lawyers who participated in their studies are likely to be relevant to interpreting the findings of this thesis. However, their findings are necessarily limited as far as providing a more complete picture of present prospects. While it is anticipated, based on their work and similar studies, that legal experts will indeed be shown to possess different cognitive traits from their less expert colleagues practising in the same specialist field of law, how readily they can be identified and measured remain open questions.

[135] Ibid.

The next chapter examines in detail the methodological elements that underlie this thesis. These include references to many of the above-noted theories and research, which this study seeks to extend within a single specialist legal-knowledge domain.

III METHODOLOGY

The methodology used in this thesis relies primarily on the approaches and techniques used in the studies reviewed in the previous chapter. That review began with an investigation into how lawyers think and ended with an analysis of studies that have looked specifically at the cognitive aspects of legal expertise. The conclusion of this review was that previous researchers in this area have predominantly focused on generic legal thinking skills and novice-expert comparisons, and as a consequence there currently exists no adequate response to this thesis' research question:

> In what readily identifiable and measurable ways do legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts?

To answer this question, the methodology developed for the present study involved the following five stages or steps: 1. Establishing a theoretical framework for analysing cognitive differences between participants as they engage in representative tasks within their domain of specialist legal expertise; 2. Securing the co-operation of a sufficient number of suitably-qualified legal specialists from law firms, economics consultancies and government agencies across the major business centres of Australia; 3. Ranking these specialists according to their levels of likely expertise in competition law; 4. Identifying performance and behavioural differences between these participants reflecting or associated with their different levels of likely expertise; and, 5. Inferring from those differences underlying cognitive processes that distinguish legal specialists according to their levels of domain expertise, rather than their levels of technical legal knowledge.

Part A of this chapter describes the theoretical underpinnings of the think-aloud verbal-protocol analysis methodology developed for this study. First, the advantages and concerns with think-aloud problem solving as a method of knowledge elicitation and the assumptions that underlie the analysis of resultant concurrent vocalizations are discussed. This is followed by a discussion of the choice of representative task and why merger review cases were used to facilitate participants' engagement in legal risk assessments requiring their specialist competition law skills.

Part B describes the process for selecting participants for this study. This includes a discussion of the threshold requirements chosen to ensure that all participants would be above the level of novice in the same specialist field of law while also allowing for a range of expertise levels extending from apprentices to masters within the meanings of these terms under Hoffman's Scheme of progressive expertise.[136] The details of how prospective participants were identified are also explained, as is the method used to solicit their involvement on a voluntary basis.

Part C describes the theoretical basis for ranking participants according to their levels of likely expertise. This includes a review of previous studies involving lawyers and more generally. These studies provide guidance on the elements of an effective and practicable ranking methodology. This part also includes a discussion of the five measures of likely expertise that were ultimately chosen, namely, promotion to partnership or a position with similar sign-off responsibility, the possession of over 10,000 hours or 10 years' experience in the field of competition law, evidence of conceptual ability, engagement in exceptional reasoning strategies, and examples of material comprehension errors.

Part D outlines the procedures followed during the testing process used in this study. This includes an overview of the think-aloud problem solving instructions given to participants prior to their analysing the test cases and an explanation of the method by which the test cases themselves were selected and presented to participants via a standardized on-line testing platform. Other technical issues relating to how the interviews were conducted by long-distance telephone calls and how transcripts were created and analysed using dedicated software solutions are also described.

Part E begins with an explanation of how the verbal protocols of participants were assessed and compared on an inter-group basis to identify cognitive differences associated with different levels of likely expertise. It then discusses the broad, exploratory approach adopted for the analysis of verbal data generated during this study. This is followed by an introduction to the three-stage problem solving framework used to categorise and track the cognitive tasks on which participants focused when assessing legal risk, and the choice to subsequently investigate aspects of

[136] Hoffman's Scheme of progressive expertise was introduced in the previous chapter as an alternative to knowledge-based frameworks for distinguishing between different levels of expertise. For a description of the characteristics of individuals at each level of this framework refer to Table 2.2 above.

participants' intuitive and analytical reasoning. General and specific expertise traits previously identified in past studies are also listed as analytical references.

Part F summarises the key methodological choices made in this study in service of generating an empirically-based response to the research question.

A *Think-Aloud Verbal Protocol Analysis*

Think-aloud verbal protocols have been used to investigate the cognitive processes of experts across a wide range of human activities.[137] Studies using this methodology were undertaken as early as the 1920s.[138] But it was Ericsson and Simon's 1984 book, *Protocol Analysis: Verbal Reports as Data* ('Protocol Analysis'), which has to date provided the most authoritative discussion of this analytical approach.[139] That text, which was republished in 1993, is still routinely cited as a foundational source of methodological theory in think-aloud studies (including in each of the previous think-aloud verbal protocol studies involving lawyers discussed in the previous chapter), and this thesis adopts the same approach.

Crandall, Klein and Hoffman describe think-aloud problem solving as a simple methodology with application to almost any task.[140] It involves participants being instructed 'to speak their thoughts as they work on specially chosen problems, and do so as if they are "speaking to themselves".'[141] It is important that participants do not try to explain what they are doing explicitly, but rather focus on solving the problem at hand. Explanations risk altering a participant's normal cognitive processes insofar as they must reflect on and reconfigure their thinking in order to verbalise an explanation.

---

[137] A recent study identified 1,926 studies utilizing 'think aloud' problem-solving and protocol analysis methodology in areas as diverse as predicting sports results, playing chess, weather forecasting, learning languages, investment portfolio selection, consumer choice, software design, clinical medicine, memorizing restaurant orders, reading ability, human aging, narrative writing and board games. See Mark C Fox, K Anders Ericsson and Ryan Best, 'Do Procedures for Verbal Reporting of Thinking Have to Be Reactive? A Meta-Analysis and Recommendations for Best Reporting Methods,' (2011) 137 *Psychological Bulletin* 2, 316. Think-aloud verbal protocol analysis has been used successfully in a number of previous studies of legal thinking skills, as discussed in the previous chapter.

[138] See, for example, J B Watson, 'Is Thinking Merely the Action of Language Mechanisms?' (1920) 11 *British Journal of Psychology* 87 and K Duncker, 'A Qualitative (Experimental and Theoretical) Study of Productive Thinking (Solving Comprehensible Problems)' (1926) 33 *Pedagogical Seminary* 642.

[139] Anders K Ericsson and Herbert Simon, *Protocol Analysis: Verbal Reports as Data* (The MIT Press, 1984/1993).

[140] Beth Crandall, Gary Klein and Robert R Hoffman, *Working Minds: A Practitioner's Guide to Cognitive Task Analysis* (The MIT Press, 2006) 95.

[141] Ibid.

Think-aloud problem solving overcomes a fundamental difficulty in the study of expert problem-solving, namely, that experts can find it difficult if not impossible to explain their thought processes to researchers in such a way that facilitates direct cognitive analysis. Lundeberg, for instance, in her study on legal reading skills explicitly acknowledged that direct questioning of experts was unlikely to give her the information she needed because experts accustomed to performing tasks within their areas of expertise are often unaware of the specifics of what they do, let alone how they do them.[142]

This so-called 'paradox of expertise'[143] connotes both the existence of tacit knowledge[144] and a lack of self-awareness on the part of the expert, who may never have thought to explain the details of their actions, either because the mental steps they perform happen automatically (and therefore defy introspection) or because the information seemed to the expert to be obvious and therefore did not warrant explication. In this context Ericsson and Simon contend that the use of their methodology in expertise studies 'is the most effective way to elicit knowledge about [cognitive] methods.'[145]

The basic assumptions underlying think-aloud problem-solving are not controversial, although some theoretical concerns have been expressed. One concern is that the act of vocalisation may change the way that a participant naturally thinks, which is known as reactivity.[146] This issue has been addressed directly in a number of studies, one of the most recent being Fox, Ericsson and Best's 2011 meta-study of almost 2,000 studies in which reactivity was shown not to have been a confounding issue.[147] As long as the

---

[142] Mary A Lundeberg, 'Metacognitive Aspects of Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407, 409.

[143] This phrase is discussed directly by Christensen, although as she explains the concept of the paradox of expertise has long been acknowledged. See Christensen, Leah M, 'The Paradox of Legal Expertise: A Study of Experts and Novices Reading the Law' (2008) *Brigham Young University Education and Law Journal* 53.

[144] The concept of tacit knowledge, as noted previously, is associated with the work of Polyani, who argued that knowledge has personal and tacit elements which cannot be made explicit, and that as a consequence knowing how to do something involves far more than simply the knowledge of the required actions. Of direct relevance to the present research is his often quoted statement that 'we can know more than we can tell.' Michael Polyani, *The Tacit Dimension* (Doubleday, 1966) 4.

[145] Ericsson and Simon, *Protocol Analysis*, above n 139, xli.

[146] For a discussion of the issue of reactivity in think-aloud problem-solving studies, see Mark C Fox, K Anders Ericsson and Ryan Best, 'Do Procedures for Verbal Reporting of Thinking Have to Be Reactive? A Meta-Analysis and Recommendations for Best Reporting Methods,' (2011) 137 *Psychological Bulletin* 2.

[147] Ibid 316. See also the discussion of this issue and elaborations on the underlying assumptions in Ericsson and Simon, *Protocol Analysis*, above n 139, where the authors cite over 30 relevant studies.

established procedures for conducting this method of verbal data elicitation are strictly followed, there should be no material concerns of this kind.

Ericsson asserts that, 'the closest connection between actual thoughts and verbal reports is found when people verbalize their thoughts that are spontaneously attended during task completion.'[148] He also reiterates, based on data analysed from a large number of empirical studies, think-aloud problem solving can reveal a study participant's 'thoughts in a manner that does not alter the sequence and content of the thoughts mediating the completion of a task and therefore should reflect immediately available information during thinking.'[149]

The think-aloud problem solving procedures described in Protocol Analysis and followed in this study as described in Part D below, are designed to ensure that the verbalization of a participant's thoughts is either Level 1 or Level 2 verbalization, to use Ericsson and Simon's terminology.[150] Level 1 describes verbalization that is 'simply the vocalization of covert articulatory or oral encodings,' without any intermediate processing such as happens when effort is expended for the purpose of communicating with other people.[151] Level 2 verbalization involves the explication of what a participant is thinking about. This requires additional processing (and therefore can slow down the participant's response time), but does not involve the participant explaining what they are thinking (which is Level 3 verbalization).[152] This is why the instructions and prompts given to participants in this study made clear that they were not to explain what they were doing, but simply to 'speak to themselves.'[153]

Levels 1 and 2 are the types of verbalizations that studies such as the present one seek to elicit and analyse. The 'stream of consciousness' aspect of these levels distinguishes them from Level 3's introspective dimension, in which an individual reflects on and reorganises their thoughts, often in ways that are considered more socially or professionally acceptable. The elicitation techniques used in think-aloud problem

[148] K Anders Ericsson, 'Protocol Analysis and Expert Thought: Concurrent Verbalizations of Thinking during Experts' Performance on Representative Tasks' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 223, 227.
[149] Ibid.
[150] See Chapter 2 of Ericsson and Herbert Simon, *Protocol Analysis*, above n 139.
[151] Ibid 79.
[152] Ibid.
[153] See further the Methodology Page from the study website reproduced below in Figure 3.1.

solving tests are intended to prevent participants entering into this third level of verbalization, which is least likely to be reflective of spontaneous, mediating thoughts.

There remain two further concerns regarding the completeness of think-aloud verbalisations postulated by Ericsson and Simon. The first reflects Duncker's observation that while the knowledge elicited by this means may be reliable, it may be that what it does not (and cannot) contain is just as important.[154] Such omissions could lead to an incomplete and possibly erroneous understating of cognitive activity since only those verbalizations that are recorded can be counted. However, this problem is likely to be less acute when reliance on the comprehensiveness of protocols is reduced, such as in this study where patterns of thinking behaviour are analysed rather than measures that assume all cognitive processes have been revealed.[155]

The other concern is that some automated responses may occur without the participant being aware of them, let alone verbalizing these cognitive events. As Ericsson and Simon have acknowledged, 'many highly overlearned processes operate automatically without leaving any more trace than their final result in [short term memory].'[156] To safeguard against such events diminishing the value of transcribed verbalisations in think-aloud studies, Middleton recommends using novel and challenging problem-solving scenarios.[157] This approach has its basis in Chi et al's contention that in such situations there will necessarily be reduced reliance on automatic responses.[158] The present study, like that undertaken by Chay who also followed Middleton's suggestions, sought to use 'authentic' and ecologically valid scenarios that were familiar to participants in terms of task procedure and required outcome, but which were novel in a substantive sense given that none of the test cases had been previously

[154] K Duncker, *On Problem Solving* (Greenwood Press, 1972 (L S Lees Trans.)) (Original work published 1945) 11.
[155] For instance, a quantitative study seeking to measure and compare the number of questions study participants posed to themselves as they solved problems would need to assume that all questions by all participants were verbalized during the relevant problem-solving tests. If some participants asked themselves questions, but did not verbalise that cognitive event (or did so indirectly by stating a response seemingly out of thin air), comparing the number of times questions were vocalised by participants would not necessarily reflect the number of questions they may have posed during the testing process. A qualitative assessment of the substantive issues raised, however, could avoid this acute problem.
[156] Ericsson and Simon, *Protocol Analysis,* above n 139, 126-127.
[157] H E Middleton, *The Role of Visual Mental Imagery in Solving Complex Problems in Design* (PhD Thesis, Griffith University, 1998).
[158] M T H Chi, P Feltovich and R Glaser, 'Categorization and Representation of Physics Problems by Experts and Novices' (1981) 5 *Cognitive Science* 121.

considered by any participant and all contained unique facts requiring individualised analysis.[159]

The remainder this part is divided into three sections. The first describes the central importance of designing a representative task to test the essence of expertise within the relevant knowledge domain. The second explains why legal-risk assessments in the context of merger review cases were considered to be appropriately unique to the practice of competition law, which was the relevant knowledge domain chosen for this study. The third section argues that such assessments involve cognitive tasks that only competition law specialists are trained to undertake, thereby ensuring that the observed cognitive differences between participants with different levels of specialist expertise would relate to these specialist skills and not to generic legal skills that lawyers in other areas of law may possess.[160]

The overriding objective in each of these contexts – and in the overall design of this study – was to ensure that only differences in participants' levels of specialist legal expertise would be subjected to cognitive analysis, and not simply the amount of technical legal knowledge they possessed or how well-developed their generic legal skills were, which would be unavoidable if the choice of a representative task merely tested participants' reading skills or their ability to interpret court cases outside their field of legal expertise. At the same time, the selected representative task needed to be capable of being performed by all participants possessing the relevant specialist legal knowledge. Accordingly, the problem-solving task had to be representative of a standard area of practice even amongst relatively junior competition law specialists.

1 *Representative Task*

The choice of an appropriate representative task is critical when developing a study based on think-aloud verbal protocol analysis. Ericsson describes the process as beginning with the identification of 'those naturally occurring activities that correspond to the essence of expertise in a domain' and the rejection of those tasks that would merely reveal the cognitive differences between novices and experts in a specialist

[159] Allan James Chay, *Lawyer Problem Solving: An Investigation of the Knowledge Used in Solving Practical Legal Problems* (PhD Thesis, Griffith University, 2006) 75.
[160] This reflects the ultimate aim of testing for competition law expertise, rather than legal expertise more generally or the possession of technical legal knowledge by only some participants, whereby the relevant tests do not test cognitive skills but rather just technical knowledge.

knowledge domain.[161] It is also important to ensure that the conditions in which experts are tested are carefully planned and executed to ensure the preservation of the highest possible methodological reliability. Where this is achieved and conditions of reproducibility are met, researchers can begin 'to examine the specific mediating mechanisms with experiments and process-tracing techniques … [and] measure and compare the performance of less-skilled individuals on the same tasks.'[162]

This study focused on how specialist competition lawyers and economists think (in terms of assessing legal risk), which was assumed to be different from how other lawyers in other specialist fields think – and from how lawyers displaying merely generic legal skills would approach the same competition law risk-assessment task.[163] Because only competition law specialists possess the necessary sub-domain knowledge required to competently undertake such a task, they necessarily think differently from other lawyers and economists who do not possess this knowledge.

Competition law specialists may still use the same generic cognitive processes as other legal specialists in terms of the basic building blocks of their analytical reasoning. However, knowledge and thought are intertwined at the applied level where the most capable intellectual property lawyer, for instance, would likely struggle with even a straightforward legal-risk assessment task in competition law. As Mitchell and Weinstein have demonstrated empirically,[164] a lack of domain knowledge can turn legal experts into virtual novices when it comes to applied legal reasoning in an unfamiliar field of law.

---

[161] Ericsson, above n 148, 231. Ericsson further cautions, 'It is important to avoid the temptation to study differences in performance between experts and novices because there are readily available tasks to measure such differences.'

[162] K Anders Ericsson, 'An Introduction to Cambridge Handbook of Expertise and Expert Performance: Its Development, Organization, and Content,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 3, 13.

[163] This conceptual approach comports with Okamoto's contention that thinking like a 'deal lawyer,' which he claimed requires an understanding of business and financial mathematics, is 'something very different from what we have traditionally meant by "thinking like a lawyer".' Karl S Okamoto, 'Teaching Transactional Lawyering' (2009) 1 *Drexel Law Review* 69, 124. In the present context, competition lawyers think more like economists, which is ostensibly also something very different from the traditional idea of how a lawyer thinks.

[164] John B Mitchell, 'Current Theories on Expert and Novice Thinking: A Full Faculty Considers the Implications for Legal Education' (1989) 39 *Journal of Legal Education* 275; Ian Weinstein, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1.

A defining feature of competition law practice is the routine application of economic analysis.[165] In this regard, the International Competition Network ('ICN') states in its investigative handbook, 'The practice of Antitrust [competition law] is in essence both a legal and an economic exercise. This truism is widely recognized, and, indeed, in most countries, legal and economic experts combine to analyze antitrust issues.'[166] Senior competition law practitioners have similarly been quoted as stating that competition law 'is all about economic analysis.'[167]

This is not to confuse the present research with being concerned with the economic analysis of law, which has developed as its own area of scholarship.[168] Rather, competition law is a form of 'economic law' that derives from statutory rules intended to enhance community welfare through the promotion of competition and economic efficiency, as well as related industrial policies.[169]

Alternative names for this area of legal practice are antitrust and restrictive trade practices law. These names are favoured in different jurisdictions, but refer to essentially the same thing. In this study, the competition law descriptor is used.

The practice of competition law also involves issues of statutory interpretation, keeping up with the latest case-law developments and a range of other generic legal and administrative tasks familiar to practitioners in this field. However, these aspects of what competition law specialists do are not of direct interest in the present context,

---

[165] As Elhauge states: '… [I]n competition law the legal doctrines often consist of vague formulations, like "dominant position", "monopoly power", "abuse of dominance", or "exclusionary conduct", that are devoid of real content unless one understands the underlying analytical model and economics.' Einer Elhauge, 'How Should Competition Law Be Taught?' (2008) 4(1) *Competition Policy International* 267, 269. See further E Elhauge and D Geradin, *Global Antitrust Law & Economics* (Foundation Press, 2007).

[166] International Competition Network, *ICN Investigative Techniques Handbook for Merger Review* (June 2005) 53.

[167] Reported statement of Peter Freeman QC, chairperson designate of the UK Competition Appeals Tribunal with 25-years' experience as a specialist competition lawyer: Sam Chadderton, 'Time for Competition Law to "Get Real," Says Incoming CAT Chairman,' *The Lawyer*, 11 January 2013.

[168] For a brief overview of the field of law and economics see Rubin, Paul H, *The Concise Encyclopedia of Economics* at http://www.econlib.org/library/Enc/LawandEconomics.html accessed 16 April 2012.

[169] For example, section 2 of the Australian *Competition and Consumer Act* 2010 states, 'The object of this Act is to enhance the welfare of Australians through the promotion of competition and fair trading and provision for consumer protection.' Hong Kong, a jurisdiction which has one of the newest competition law regimes in the world, is promoting its new competition rules to reflect 'the Government's competition policy … to enhance economic efficiency and free flow of trade, thereby benefiting consumer welfare,' Research and Library Services Division of the Hong Kong Legislative Council Secretariat, *Competition Policies in Selected Jurisdictions*, 25 June 2010, paragraph 1.1.2.

although they could legitimately be – and, as was shown in the previous chapter, have been – the focus of empirical studies of how lawyers think in more general terms.

The essence of a competition law specialist's expertise for present purposes is in their performance of applied legal and economic analysis within a specific regulatory and policy setting. This combined analysis has, to the researcher's knowledge, no analogue in other areas of law and for this reason constitutes an appropriate focus for the study of competition-law expertise.

2 *Merger Review Cases*

Merger review cases are a common feature of many, and ostensibly most, competition law specialists' practices.[170] A review of the practice profiles of those Australian, New Zealand, UK and US law firms which maintain specialist competition law departments and are recognised as leaders in this field in legal directories published by *Chambers and Partners*,[171] *Who's Who Legal*[172] and *The Legal 500*,[173] all promote on their websites their expertise in handling merger review cases as a core – and often as their lead – service offering.[174] In this regard, all participants in this study confirmed that they had had direct involvement in a number of merger review cases during their careers, including the least experienced. Moreover, 85% of them volunteered that merger review cases were a focus of their professional practice.[175]

[170] In this paper, the term merger is used to include acquisitions of shares or assets as defined by relevant merger control laws.
[171] http://www.chambersandpartners.com/.
[172] http://whoswholegal.com/.
[173] http://www.legal500.com/.
[174] For example, in Australia the law firm Allens (www.allens.com.au) describes its competition law team as offering its clients 'experience, expertise and commonsense in dealing with the full spectrum of trade practices issues, including: a) mergers and acquisitions; b) joint ventures; anti-competitive conduct allegations …' In New Zealand, law firm Chapman Tripp (www.chapmantripp.com) begins the description of its competition law expertise with the claim that its competition law experts 'are known for their ability to succeed in mergers and acquisition, often in complex and changing markets.' In the UK (and globally), law firm Freshfields Bruckhaus Deringer (www.freshfields.com) states that 'Clients come to us for advice on the complete range of antitrust/competition, regulatory and trade issues spanning merger control and joint ventures; restrictive practices; market dominance …' In the US (and globally), law firm Cleary Gottlieb (www.cgsh.com) states on its Antitrust and Competition page, 'Our lawyers respond to clients' needs in a broad range of antitrust/competition law matters, including: a) Pre-Merger notification and Merger Control issues; b) Antitrust audits and compliance programs/cartel investigations; c) Monopolization/dominance cases …'
[175] This was in response to the following question posed in the pre-task questionnaire (See Appendix A): 'What kinds or types of competition law matters (a) do you mainly work on at present, and (b) have you mainly worked on in the past?' The remaining 15% of participants stated that their current involvement in substantial litigation matters had restricted their mergers work for the time being. This confirmed the researcher's initial view that merger review cases are not only the exclusive province of competition

Merger review cases are concerned with future reductions in competitive tensions within specific markets as a direct result of, for instance, two vigorous competitors combining their businesses or a dominant business taking over a smaller, maverick-type business which was previously the instigator of price-wars in a particular industry. The concern in these situations is that reductions in competitive rivalry as a result of a proposed merger could lead to less efficient market structures, higher prices and reduced choice for consumers.[176]

Procedurally, the merger review process centres on the role of a competition authority, commonly an independent statutory or government body, which assesses whether or not a merger is likely to violate a statutory prohibition against anti-competitive mergers as found in most competition laws. According to the ICN – which is an association of national and multinational competition authorities representing the world's largest competition law jurisdictions[177] – merger review cases involve competition authority staff attempting 'to predict a merger's competitive impact to prevent any competitive problems before they materialize,' while at the same time seeking to ensure procedural consistency, predictability and transparency.[178]

Based on a multi-year study by the ICN, such prohibitions exist in almost every competition law statute in every competition law jurisdiction,[179] and Clarke has recorded that an estimated 110 countries have statutory regimes of this kind.[180]

---

lawyers and competition economists owing to their highly specialist substantive and procedural natures, but are also a central feature of a competition law specialist's practice.

[176] The Australian Competition and Consumer Commission describes the significance of mergers in a competition sense in the following way: 'In the vast majority of mergers, sufficient competitive tension remains after the merger to ensure that consumers and suppliers are no worse off. Indeed, in many cases, consumers or suppliers benefit from mergers. In some cases, however, by altering the structure of markets and the incentives for firms to compete, mergers have anti-competitive effects.' Australian Competition and Consumer Commission, *Informal Merger Review Process Guidelines* (Australian Competition and Consumer Commission, September 2013) 5. See also: Richard Whish and David Bailey, *Competition Law* (Oxford University Press, 7th Edition 2012); Ulrich Schwalbe and Daniel Zimmer, *Law and Economics in European Merger Control* (Oxford University Press, 2009).

[177] The ICN or International Competition Network was launched on 25 October 2001 by 'top antitrust [competition law] officials from 14 jurisdictions – Australia, Canada, European Union, France, Germany, Israel, Italy, Japan, Mexico, South Africa, United Kingdom, United States, and Zambia.' See http://www.internationalcompetitionnetwork.org/about/history.aspx

[178] International Competition Network, *ICN Recommended Practices for Merger Analysis* (International Competition Network, 2008) 3.

[179] According to relevant lists curated by the International Competition Network's Mergers Working Group, over 60 jurisdictions have legislated merger control laws, including Mexico, Kenya, Australia, Russia, European Union, Jordan, Iceland, Singapore, Cyprus, United States, Panama, South Africa, Estonia, Canada, Germany, Turkey, Korea, Costa Rica, Ireland, Barbados, Israel, Japan and Zambia. See http://www.internationalcompetitionnetwork.org/

In some countries, merger proponents (usually when involved in transactions exceeding certain combined-turnover amount or other thresholds) may be required to notify their merger proposals to the relevant competition officials.[181] In other jurisdictions, notification is voluntary, but informal procedures exist whereby officials can review proposed transactions either confidentially or publicly, and then provide their views as to any legal compliance issues.[182] In cases where there are no such issues, a competition authority may provide a qualified letter of comfort to merger proponents on a pre-assessment basis without undertaking a public review.

In Australia, as one of the two relevant jurisdictions in this study, most publicly reviewed merger transactions are cleared without official comment, while a small percentage are cleared with conditions attached. An even smaller number are opposed by the competition authority on grounds of likely anti-competitive detriment.[183]

A voluntary notification framework and an informal, public review process were the general background conditions with which the participants in this study – who were specialists in either Australian or New Zealand competition law – were most familiar. Accordingly, this was the regulatory framework of interest in this study.

3 *'The Essence' of Competition Law Expertise*

The representative task in this study required participants to provide legal-risk assessments relating to merger transactions under review by the Australian Competition and Consumer Commission ('ACCC') or the New Zealand Commerce Commission ('CC'). Rather than provide a definitive view (which was impracticable given the limited information provided in the test-case documentation), the objective was for each participant to provide the kind of advice merger proponents typically seek when

---

[180] Julie Clarke, *The International Regulation of Transnational Mergers* (PhD Thesis, Queensland University of Technology, 2010) 12.

[181] This is the case in, for example, Canada (see Part IX of the Competition Act), Israel (see Chapter III of the *Restrictive Trade Practices Act* 5748-1988), and the United States (see section 7A of the *Clayton Act*, 15 USC, the *Hart-Scott-Rodino Improvements Act of 1976* and the *Premerger Notification Rules*, 16 CFR Parts 801-803 (2008)).

[182] This is the case in, for example, Australia (see section 50 of the *Competition and Consumer Act* 2010), New Zealand (see section 47 of the *Commerce Act* 1986) and the United Kingdom (see the *Enterprise Act* 2002).

[183] Based on statistics published by the Australian Competition and Consumer Commission in its annual reports for 2009/10, 2010/11 and 2011/12, the total number of mergers reviewed confidentially and non-confidentially was 340, 377 and 321, respectively. Of these, 1, 3 and 8 were publicly opposed, 6, 4 and 6 were opposed (or significant concerns were raised) confidentially. The total percentages of all mergers unconditionally cleared (including those cleared in pre-assessment decisions) were 91%, 95% and 88%, respectively.

assessing the ease or difficulty of securing the necessary approvals from a responsible competition authority.[184]

Regarding the methodological efficacy of using merger cases to test the essence of competition law expertise, most of a competition law specialist's time in a merger review case is spent, in a substantive sense, undertaking an applied legal and economic analysis. They must also deal with procedural and tactical issues. But their ability to understand and master the economic implications of a proposed transaction within the parameters of the relevant merger laws, arguably dictates to a significant extent all other aspects of their brief.

The amount of black letter law that needs to be navigated in most merger cases in virtually any competition law jurisdiction can be reduced to one simple question: Is this merger transaction likely to materially harm future competition?[185] There is seldom if ever a need to inquire into issues such as intent or who-said-what-to-whom as happens in price fixing, monopolization and many other restrictive trade practices cases typically handled by competition law specialists.[186] Statutory interpretation and forms of generic legal analysis are therefore less important, in relative terms, in merger review cases as compared to cases involving other competition law prohibitions.

In most merger matters, the necessary economic analysis begins with the market definition process, which Beaton-Wells – quoting from the landmark Australian case *Re QCMA*[187] – describes as 'the "essential first step" in the assessment of present competition and likely competitive effects.'[188] According to Round, market definition is a well-established analytical exercise which is governed by principles that are

[184] Such approvals are commonly stated as conditions precedent in the transaction documentation, thereby making the risk assessment of competition law advisers an important input into the deal-making process.
[185] Under Australian and New Zealand competition laws, which are directly relevant to this study, the respective prohibitions apply to mergers or acquisitions of shares or assets that 'would have the effect, or be likely to have the effect, of substantially lessening competition in a market' (Section 50 of the *Australian Competition and Consumer Act* 2010), or 'would have, or would be likely to have, the effect of substantially lessening competition in a market' (Section 47 of the *Commerce Act* 1986). Section 4G of the Australian law and section 3(2) of the New Zealand law further state that 'references to the lessening of competition' in this context include references to preventing or hindering competition.
[186] Of course, there can arise questions such as whether a transaction qualifies as a merger within the meaning of the law, which is a technical legal issue. But it is neither possible nor efficacious to try to investigate every type of issue a competition law specialist must deal with in a study such as this, where it is the essence of the domain expertise (ie, applied economic analysis) that is of primary interest.
[187] *Re Queensland Co-Operative Milling Association Ltd* (1976) ATPR 40-012.
[188] Caron Beaton-Wells, *Proof of Antitrust Markets in Australia* (The Federation Press, 2003) 10.

generally settled.[189] Also relevant to the present study, which involves both Australian and New Zealand competition law specialists, are the authoritative statements of Smith and Walker,[190] and Brunt,[191] who as practitioners and adjudicators in both jurisdictions claim that the same basic approach to market definition is followed in both Australia and New Zealand.

Importantly, however, the market definition process is not limited only to merger cases. It is also an essential element in the application of every competition law prohibition concerned with market power issues, such as unilateral abuses of market dominance and forms of exclusionary conduct. Moreover, it is integral to the analysis of barriers to entry and other issues relevant to the contestability and competitiveness of markets. Insofar as it is central both to the work of a competition law specialist and to the analysis of merger review cases, it is an ideal focus for a representative task of the kind required for a study such as this.

In the context of competition law in general, Beaton-Wells characterises market definition as 'an evaluative exercise involving mental processes akin to those involved in dealing with standards derived from the common law or equity such as reasonableness, fairness, just cause or hardship.'[192] This view, which is supported by Gault's contention that issues of substitution in market analysis require 'informed assessment' in a manner akin to judicial assessments of negligence, reasonableness and fairness,[193] highlights the importance of judgment, wisdom – and hence expertise – in the assessment of merger review cases where market definition is a central task.

These comments by Beaton-Wells and Gault also underline the deeper significance of choosing a representative task that is solely limited to the domain of competition law but which by virtue of this very focus facilitates the discovery and analysis of cognitive skills associated with expertise across a much wider range of specialist areas of legal practice. Because only competition law specialists would be selected to undertake a

[189] D Round, 'Market Definition – A Matter of Fact, Commercial Commonsense or Economic Principles?' in A Bollard (ed), *The Economics of the Commerce Act* (New Zealand Institute of Economic Research, 1989) 20.
[190] R Smith and J Walker, 'Australian Trade Practices and the Emerging Role of "Commercial Reality" Versus Substitution in Market Definition' (1997) 5 *Competition & Consumer Law Journal* 1, 10.
[191] M Brunt, '"Market Definition" Issues in Australian and New Zealand Trade Practices Litigation' (1990) 18 *Australian Business Law Review* 86, 102.
[192] Beaton-Wells, above n 188, 83.
[193] Justice Gault, 'A Judicial Perspective on Competition Litigation,' in R Adhar (ed), *Competition Law and Policy in New Zealand* (Law Book Co., 1991) 96, 101.

representative task unique to their particular area of legal specialisation, their specialist knowledge becomes irrelevant in the sense that observed cognitive differences would exist only in terms of their levels of expertise. Issues of judgment, wisdom, insight and other expertise-related considerations could therefore be addressed without worrying about whether such differences arose merely because some participants knew more law or were more familiar with the relevant regulatory procedures than others.

As regards ancillary legal issues in merger review cases, these are for the most part issues that other lawyers might be expected to manage competently in both substantive and procedural senses insofar as they have their analogues in other areas of law. This is not the case for applied economic analysis and not for the definition of markets as described. As Elhauge notes in the context of teaching competition law skills, it is important to focus on the most distinctive elements of substantive competition law analysis, on the basis that 'procedural issues are common to many courses and can be picked up much more easily in practice. The substance of competition law [which is economic analysis] is unique …'[194]

These were important considerations from a study design perspective. Without a substantial outer-wall of legalese to penetrate in merger review cases, the need for general legal skills was significantly reduced, while at the same time key economic issues would present themselves for analysis almost immediately the task was revealed.[195] This would enable study participants with less developed legal skills – such as economists – to engage with the central problems of the task faster and more directly, while presentation of the task itself could be reduced to a few key facts without sacrificing realism or the ability of the researcher to regulate degrees of difficulty. This was considered likely to improve experimental efficiency and open the possibility of each participant completing multiple test-cases per interview session.

[194] Einer Elhauge, 'How Should Competition Law Be Taught?' above n 165, 271.

[195] With price-fixing or vertical restraint cases, as just two alternative types of competition law matters, the importance of factual details in simply understanding the basic events or arrangements at issue, means that extensive amounts of information must be absorbed before any view can be expressed. In merger transactions, the central event is instantly understood as an acquisition of shares or assets, which greatly simplifies and streamlines the initial framing of the case. The subsequent economic analysis can also be undertaken with a greater degree of abstraction from the facts, whereas price-fixing and vertical restraint cases require a heavily facts-dependent assessment just to determine whether or not the relevant laws apply, let alone what the outcome of a case is likely to be. The same applies equally to misuse of market power and abuse of dominant position cases, as well as to resale price maintenance and boycott allegations.

Further below in Part D is a discussion of the actual merger test cases used in this study and the decision to use a limited-information approach for testing purposes.

B *Participant Selection*

The next step was to develop a means of identifying prospective candidates who would constitute a sufficiently large and varied, yet focused, group of specialists all with substantial technical legal knowledge of competition law. It was critical to ensure as far as practicable that there would be participants who possessed differing levels of expertise within this specialist knowledge domain, but that none of them would be considered mere novices or initiates as defined under Hoffman's Scheme. A total of between 15 to 20 participants was viewed as being both sufficient for the purposes of analysis and feasible for a single researcher to manage.[196]

It was decided that each participant should have at least five years specialist competition law experience. This baseline requirement avoided the involvement of novices or initiates, whom it was assumed would necessarily have less than 5 years specialist experience.[197] Competition law also needed to be the focus of their professional work, and not just one of a broad portfolio of areas in which they practised.[198] This restricted the pool of potential candidates to lawyers and economists working for major Australian law firms, specialist economics consultancies, or competition authorities. It also meant a focus on the main national business centres in Australia, namely, Sydney, Melbourne, Brisbane and Perth. This was because it was assumed that the work environments in smaller cities would not permit the degree of specialisation required for participants in this study.[199]

---

[196] This assessment was based on the numbers of participants involved in similar studies discussed in the previous chapter, most relevantly those of Weinstein (10 participants), Colon-Navarro (7) and Chay (4).

[197] Given the choice of legal-risk assessments using merger review cases as the representative tasks in this study, it was necessary to ensure that all participants could attempt such a task without needing to refer to relevant statutes, case law or guidelines. This was less likely to be the case with someone who had only had 'some ('minimal') exposure to the domain' or who had only 'begun introductory instruction,' which are the descriptions used for novices and initiates under Hoffman's Scheme. The basic concern was that participants with less than 5 years' experience in this specialist area or law may struggle to perform in this context. As it was, all participants who ultimately volunteered had at least 5 years' specialist experience and none indicated any substantive concerns regarding the legal issues raised by the test cases, or any concerns with any procedural issues.

[198] To ensure that these individuals were not part-time specialists who merely 'dabbled' in competition law, it was decided that each participant had to have spent at least 50% of their billable or working hours during the last twelve months on competition law matters.

[199] This assumption acknowledged the observation attributed to Adam Smith that the size of the market ultimately determines the degree to which a firm or an individual can specialize, and that as Garicano

The targeting of market-leading law and economics firms was also important insofar as they screen-out those individuals whose technical skills or personalities do not meet the requirements of these organisations. This screening function, which was formally recognised in Spence's job-signaling model where higher education provides a sorting function for would-be employers,[200] has been explicility considered by Ippolito in the context of the legal profession[201] and was also noted by Blasi.[202] In essence, those individuals with the IQ, tenacity (or, to use Duckworth et al's terminology, 'grit'[203]), social skills and technical capabilities (as well as high levels of ambition or 'passion'[204]), ultimately overtake their less-capable and less-motivated peers in the job market. These individuals become employees at the best law and economics firms which are, in turn, attuned to identifying those individuals who best match their demanding talent-requirements.[205]

It was assumed that factors such as IQ, grit and technical abilities would be consistently high across the sample of individuals in this study because they worked for these kinds of firms. This would have the effect of controlling, at least to some degree, for factors not directly associated with this study's conceptualisation of specialist competition law expertise, but which nevertheless could be important in explaining how these individuals came to be in their current positions. Competition authority employees

---

and Hubbard, and Leahey and Hunter have demonstrated, specialization is most likely to be an optimal strategy in larger cities. See Luis Garicano and Thomas N Hubbard, 'Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyer's Human Capital' (2007) 50(1) *Journal of Law and Economics* 1; Luis Garicano and Thomas N Hubbard, 'Specialization, Firms, and Markets: The Division of Labor within and between Law Firms' (2008) 25(2) *The Journal of Law, Economics, & Organization* 339.; Erin Leahey and Laura A Hunter, 'Lawyer's Lines of Work: Specialization's Role in the Income Determination Process' (2012) 90(4) *Social Forces* 1101.

[200] Michael Spence, 'Job Market Signaling' (1973) 87(3) *The Quaterly Journal of Economics* 355.

[201] Richard A Ippolito, 'The Sorting Function: Evidence from Law School' (2001) 51 *Journal of Legal Education* 533.

[202] Gary L Blasi, 'What Lawyers Know: Lawyering Expertise, Cognitive Science, and the Functions of Theory' (1995) 45(3) *Journal of Legal Education* 313, 315.

[203] A L Duckworth, T A Kirby, E Tsukayama, H Berstein, and K A Ericsson, 'Deliberate Practice Spells Success: Why Grittier Competitors Triumph at the National Spelling Bee' (2012) 2 *Social Psychology and Personality Science* 174.

[204] A Bonneville-Roussy, G L Lavigne and R J Vallerand, 'When Passion Leads to Excellence: The Case of Musicians' (2011) 39 *Psychology of Music* 123.

[205] It is also relevant to note that as their skills develop, such professionals are typically encouraged to specialise in order to maximise their productivity, as has been documented by economists such as Garicano and Hubbard, who identified the contexts in which specialization for lawyers makes the most sense (which is, in larger firms and boutique practices, and at times of high demand), and Leahey and Hunter, who have shown empirically how the more narrowly a lawyer specializes (for instance, in just one or two practice areas) the more revenue they are likely to generate for their firms and for themselves compared with pursuing a more general practising strategy. See Garicano and Hubbard, 'Managerial Leverage', above n 199; Garicano and Hubbard, 'Specialization, Firms, and Markets', above n 199; Leahey and Hunter, above n 199.

were targeted on the same basis inasmuch as demand for opportunities to work in such organisations is high owing to the quantity and quality of competition work undertaken by these 'monopolies' of competition law enforcement, and because experience with a competition authority is a well-trodden path to private-sector employment with leading law and economics firms.[206]

A balance between senior and junior participants was to be achieved by adopting a staggered selection process whereby once a number of, say, senior participants had volunteered, solicitations to more junior participants would be increased. To this end, it was assumed that those participants who were partners or at an equivalent level of seniority within their organisations (where they had sign-off responsibility) would have greater expertise than those participants who did not meet these conditions, but who were nevertheless domain specialists with at least 5 years' experience in this field. But this constituted a rough approximation only. The aim at the participant-selection stage would not be to rank participants as such, but rather to ensure a broad representative sample of participants whose cognitive skills would extend the full range of expertise possibilities.

As to making initial contact with prospective participants, it was decided to send invitational e-mails either directly and unsolicited to lawyers and economists whose profiles were listed on their organisations' websites (where their specialisation and seniority were clearly displayed), or to do so on the recommendations of other participants who permitted the researcher to use their name to solicit further prospects.

No participant would be offered payment of any kind or any other financial or non-financial incentive to participate.

An overview of the individuals who volunteered for this study is provided in the next chapter. This includes details concerning each participant's employing organisations, years of professional experience, years of specialist experience in competition law, and the number of merger matters in which they had been involved during their careers.

[206] This point was confirmed inasmuch as almost half the lawyers and economists in private practice who participated in this study had previously worked for a competition authority. In terms of Spence's Job-Signaling Model, such individuals' experience is an indicator to law and economics firms that these professionals have valuable – and rare – knowledge, skills and abilities.

C *Participant Ranking*

The third step in the methodology developed for this study was to determine a means by which participants could be ranked according to their levels of likely expertise. As no previous studies had developed such a process, at least not in a form that fitted the analytical objectives of this thesis, it was necessary to establish a reasoned and practicable basis for creating an appropriate hierarchy of legal specialists amongst and between whom cognitive performances could be compared. This meant devising a scale to measure expertise levels and a weighting system that would ensure a meaningful balance between these measures of specialist legal ability. This would be a further, and more refined, sorting of participants beyond the approximate categorisations of partners and non-partners used when soliciting volunteers. That distinction was relevant only for sampling purposes. A more detailed and systematic ranking procedure was needed to permit more fine-grained – and theoretically explicit – comparisons along the expertise continuum.

Chess players have long been popular subjects for expertise researchers in much the same way that the fruit fly has been a model organism for geneticists.[207] This is primarily because of the existence of the ELO rating scale, which has been used to rank chess players since the 1960s.[208] Gobet and Charness describe this rating scale as 'a sophisticated measurement scale for evaluating chess skill based on performance in chess tournaments.'[209] This scale has enabled expertise researchers to undertake fine-grained comparisons between the cognitive abilities of different levels of chess experts, such as those in De Groot's seminal studies comparing expert and grandmaster-level chess players[210] and in follow-up research by Chase and Simon,[211] as well as in more

[207] Fernand Gobet and Neil Charness, 'Expertise in Chess,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 523. See also, N Charness, 'The Impact of Chess Research on Cognitive Science' (1992) 54 *Psychological Research* 4.
[208] A E Elo, 'Age Changes in Master Chess Performance' (1965) 20 *Journal of Gerontology* 289 and A E Elo, *The Rating of Chessplayers, Past and Present* (Arco Chess, 2nd Edition 1986).
[209] Gobet and Charness, above n 207, 524.
[210] A D de Groot, *Thought and Choice in Chess* (Mouton, 1965).
[211] WG Chase and HA Simon, 'Perception in Chess' (1973) 4 *Cognitive Psychology* 55 and WG Chase and HA Simon, 'The Mind's Eye in Chess,' in WG Chase (ed), *Visual Information Processing* (Academic Press, 1973) 215.

general studies involving chess players by Bachmann and Oit,[212] Charness et al,[213] and Van der Mass and Wagenmakers.[214]

In other domains, the lack of a similar scale against which to rank individuals with different levels of expertise has remained a problem for expertise researchers,[215] including those involved in the cognitive analysis of legal expertise. Previous studies of cognitive differences between expert and novice lawyers (often law students), have relied on broad assumptions regarding the superior abilities and performance of the putative experts concerned.[216]

Most relevant to the present study of specialist legal expertise were the participant selection choices of Colon-Navarro and Weinstein in their think-aloud verbal protocol analyses of immigration law experts and social security disability ('SSD') claims experts, respectively. Colon-Navarro chose as his highest level of experts four

[212] T Bachmann and M Oit, 'Stroop-like Interference in Chess Players' Imagery: An Unexplored Possibility to be Revealed by the Adapted Moving-Spot Task' (1992) 54 *Psychological Research* 27.
[213] N Charness, M Tuffiash and T Jastrzembski, 'Motivation, Emotion, and Expert Skill Acquisition,' in D Dai and RJ Sternberg (eds), *Motivation, Emotion and Cognition: Integrative Perspectives* (Erlbaum, 2004) 299.
[214] HLJ Van der Maas and E-J Wagenmakers, 'A Psychometric Analysis of Chess Expertise' (2005) 118 *American Journal of Psychology* 29.
[215] Gobet and Charness, above n 207, 524.
[216] The first recorded use of think-aloud protocol analysis involving experts in a legal context was Lundeberg's study in which legal experts were identified solely by their being either lawyers or law professors with at least 24 months' practising or teaching experience (Mary A Lundeberg, 'Metacognitive Aspects of Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407). Mitchell's research subjects were fellow academics from the same law faculty in which he taught. He did not attempt a formal think-aloud study, however, and neither did he identify the level of expertise of the individuals involved, other than to note that some had criminal law expertise while others did not (Mitchell, above n 164, 280). Marchant et al, who used tax advisors rather than tax lawyers in their study, identified their experts as tax professionals who had one to eight years' experience as tax professionals and who were employed by the same multinational public accounting firm (Garry Marchant et al, 'Analogical Transfer and Expertise in Legal Reasoning' (1991) 48 *Organizational Behaviour and Human Decision Processes* 272, 274). Nievelstein et al compared performance against a sorting task involving 48 first and second-year law students and 12 academics who 'specialised in civil law.' Their expert group had, on average, '5.9 years of professional experience after obtaining their PhD' (Fleurie Nievelstein et al, 'Expertise-Related Differences in Conceptual and Ontological Knowledge in the Legal Domain' (2008) 20(6) *European Journal of Cognitive Psychology* 1043, 1048) However, the authors did not disclose the variation in experience amongst these experts nor their specific areas of civil law expertise. The same participants were the subjects of Nievelstein et al's subsequent journal articles concerning information requirements for novices and experts performing legal problem-solving tasks. See: Fleurie Nievelstein, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Effects of Conceptual Knowledge and Availability of Information Sources on Law Student's Legal Reasoning' (2010) 38 *Instructional Science* 23; Fleurie Nievelstein, Tamara van Gog, Gijs van Dijk and Henny P A Boshuizen, 'Instructional Support for Novice Law Students: Reducing Search Processes and Explaining Concepts in Cases' (2011) 25 *Applied Cognitive Psychology* 408.

experienced lawyers from Boston.[217] His assumption was that lawyers who had specialized in immigration law for between three and 14 years and who had handled between 500 to 1,500 immigration problems[218] could be considered experts vis a vis two 'experienced novices' who were law students with both theoretical training and some clinical experience, and vis a vis one inexperienced novice with only theoretical training in immigration law.[219]

In Weinstein's study, three of his 10 participants were designated experts on the basis of their being 'outstanding practitioners,' but their credentials were not stated.[220] The author simply indicated that their expertise arose from their experience dealing with SSD claims of the general type chosen for the think-aloud problem solving task in the study.

Chay's more recent doctoral research, which also involved verbal protocol analysis involving legal specialists, used two experienced family law specialists and two law students.[221] His first designated expert was a lawyer who specialized exclusively in family law matters, but who had practiced in the area for only 6 years. The second was a lawyer with 22 years of experience in family law matters, but who also practiced concurrently in the fields of criminal and business law. The two law students in Chay's study were law school graduates who had yet to complete their practical legal training qualifications.

In the absence of a formal scoring system like that used to rank chess players, and recognizing the methodological shortcomings of previous researchers' assumptions concerning their choice of legal experts, the approach adopted in this study utilized multiple measures of expertise to create a ranking of participants according to their likely levels of specialist legal expertise. Which is to say, no attempt was made to determine levels of expertise in an absolute sense, but rather to identify which participants were more likely (and which were less likely) to be experts according to a set of five measures of expertise informed, for the most part, by previous research.

---

[217] Fernando Colon-Navarro, 'Thinking Like a Lawyer: Expert-Novice Differences in Simulated Client Interviews' (1997) 21 *The Journal of the Legal Profession* 107, 121.
[218] Ibid.
[219] Ibid.
[220] Weinstein, above n 164, 18.
[221] Chay, above n 159, 82.

This approach comports with Chi's review of the so-called 'relative expertise' approach to expertise research, which obviates the need to define expertise in absolute terms.[222] The theoretical construct underlying this approach was that of a continuum of expertise, which Chi argues is better for identifying the specific skills and knowledge that less expert individuals need to progress to higher levels within a particular knowledge domain.[223]

Given the focus in this study on higher-levels of specialist expertise rather than on simply experts versus novices, an overlay of probabilities acknowledged that it was not possible to discover all relevant facts about individual participants such that a definitive view of their position on the continuum of expertise might be determined. For instance, certain expert behaviours may not manifest in ways that that would be recognized or recorded within the chosen analytical framework. This acknowledges Duncker's concern that the analysis of verbal protocols is necessarily limited to only that information which is verbalized.[224]

It was therefore decided to conceptualise the rankings process in terms of 'more likely to be an expert' and 'less likely to be an expert.' This was to make clear that higher-ranked participants were simply those who had had their positive performances against the chosen measures recorded and assessed. Very good performance against all these measures was therefore assumed to reflect a greater likelihood that an individual participant was a higher-level expert.

Poor performance against one or more measures, however, could not be so easily associated with lesser expertise. Allowing for the limited number of measures used (and hence the unavoidably incomplete capturing of expert performance) as well as a participant's possible reliance on non-vocalized mediating protocols, the absence of good performances against the chosen measures could not be said to necessarily reflect lesser expertise. Instead, it indicated only the probability or likelihood that an individual had less expertise than another participant who was observed performing better against the same measures.

---

[222] M T H Chi, 'Two Approaches to the Study of Experts' Characteristics,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 21, 23.
[223] Ibid.
[224] Duncker, above n 154, 11.

In other words, it was important to ensure when implementing these measures that the 'absence of evidence' of such abilities was not conflated with 'evidence of absence' of those abilities, a problem associated with the so-called black swan phenomenon discussed by Taleb.[225] The chosen probabilistic approach therefore acknowledged the possibility that some participants would perform poorly against these measures in the context of this study, but may nonetheless possess high levels of relevant expertise.

Five measures of probable expertise were chosen for this aspect of the study. These were: a) Whether or not a participant had been promoted to a partnership position as a competition law specialist (which status was equated with having sign-off responsibility); b) Whether or not they had 10,000 hours (approximately 10 years) specialist experience in competition law; c) The depth of their analysis, in conceptual terms, when assessing two designated test cases; d) Their ability to generate exceptional reasoning strategies during the problem-solving tasks, indicating either high or low levels of expertise; and, e) How prone they were to making comprehension errors.

There was no assessment undertaken at this point as to whether participants offered correct or accurate legal risk assessments against the test cases. This was principally because without a complete analysis of their verbal protocols (which was considered more appropriately left until the post-test assessment), it was not possible to determine whether a seemingly accurate assessment might be the result of mere guessing or flawed reasoning. Accordingly, this measure was viewed as problematic and likely to lead to circularity problems (discussed further below) when it came to correlating assessed levels of expertise with the substantive results of the testing phase.

Moreover, correct or accurate responses to a task cannot always be associated with greater levels of expertise, and in some instances incorrect responses can be explained on the basis that the relevant individual was an expert.[226] This was observed in

[225] The black swan problem arose when western Europeans, who had only ever seen white swans, concluded that all swans were white – a statement that became a cultural truism. It was not until explorers discovered the existence of black swans in places such as Australia, that the logical error of conflating the absence of evidence of white swans with evidence that black swans do not exist was acknowledged. See Nassim N Taleb, *The Black Swan: The Impact of the Highly Improbable* (Penguin Books, 2007).

[226] As has been observed, experts do not always give the best answers and may perform worse than less expert individuals, even when operating within their areas of expertise. See Ericsson, above n 162, 4; Ericsson, K A and A C Lehman, 'Expert and Exceptional Performance: Evidence on Maximal Adaptations on Task Constraints' (1996) *Annual Review of Psychology* 47; E J Johnson, 'Expertise and

Marchant et al's study in which their assumed experts were incapable of dismissing from their minds a tax deduction rule that conflicted with the researchers' expected resolution of the test-cases.[227] Feltovich and Barrows also note how the limiting effects of contextual dependence can lead expert medical practitioners into error or at least prevent them from offering an opinion while less capable doctors freely express their views.[228] In this latter respect, it was entirely possible that some participants in the present study would refuse to give any opinion – or offer only a highly qualified response – because of insufficient information and too little time. It would have been difficult to categorise such responses as correct or accurate (or incorrect or inaccurate) in the standard sense.

With respect to the last three measures noted above, which depended on the assessment of participants' performances during their consideration of the tests cases in this study, a further concern was the circularity of ranking participants according to their levels of expertise by utilising the same test cases used to test their expertise. To the extent that such circularity might exist and influence the results of the study, it was decided that the areas of comparison subsequently chosen for protocol analysis would be unrelated (or only indirectly related) to the above three measures of cognitive development.

The chosen measures were otherwise considered consistent with the substantive tests of expertise derived from previous studies as discussed below. They also comported with Chay's contention that performance measures of cognitive development are analytically valid in a ranking context, although they can be logistically cumbersome to implement when test participants are busy professionals.[229]

Below is a more detailed discussion of the nature and relevance of all five of the above measures of likely expertise. This is followed by an explanation of the measurement

---

Decision Under Uncertainty: Performance and Process,' in M T H Chi, R Glaser and M J Farr (eds) *The Nature of Expertise* (Erlbaum, 1988) 209; R M Dawes, 'A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making' (1971) 26 *American Psychologist* 180.

[227] Mitchell, above n 164.

[228] P J Feltovich and H S Barrows, 'Issues of Generality in Medical Problem Solving,' in HG Schmidt and ML de Volder (eds) *Tutorials in Problem-Based Learning* (Van Gorcum, 1984) 128.

[229] Chay limited his selection and ranking criteria to experience-based measures, although he indicated support for and agreement with more sophisticated models that 'use the individual's cognitive development as the criteria for determining level of expertise.' His choice not to adopt these models was because 'it was not possible to use them to select research subjects without carrying out extensive investigations of each potential subject's cognitive development.' This comported with his not differentiating between several stages 'in the progress from novice to expert.' Instead, he purported to focus simply 'on any transformations of knowledge that may take place after several years of experience.' Chay, above n 159, 81.

and weighting approaches adopted for the purpose of scoring participants according to their assessed likely levels of expertise, both against each individual measure, as well as cumulatively. These measures define the ranking methodology on which this study relied for identifying associations between participants' cognitive performances and their respective levels of specialist legal expertise.

1 *Promotion to Partnership as a Competition Law Specialist*

The first measure of likely expertise was whether or not a participant possessed sign-off responsibility and, more particularly, whether their promotion to partnership or its non-law firm equivalent was based on their demonstrated expertise in competition law.

This difference, which was the basis for the partner and non-partner distinction relied on during the participant-selection process, mirrors the difference between journeymen and experts under Hoffman's Scheme where journeymen work under orders from their superiors, but experts are self-directed and can therefore be held responsible for their work in ways that journeyman are not. Accordingly, this first measure of likely expertise assumed a positive correlation between a participant's level of expertise and their ability to sign-off on their work as an authorised representative of their organisation.

This level of responsibility is typically only given to the partners of law firms (or their economist equivalents), such that only their signatures on official firm documents, in the form of letters of advice for example, are binding on their partnership or firm. Securing such responsibility generally requires the agreement of existing partners or of senior management in the case of incorporated organisations and economics firms which do not have a partnership structure. In this sense, sign-off responsibility is ostensibly a proxy for the requirement noted by Hoffman that experts must be individuals who are 'highly regarded by their peers' inasmuch as their 'judgments are uncommonly accurate and reliable.'[230] Existing partners or senior management are unlikely to bestow an ability to legally bind them and their firm on an individual who has not demonstrated these attributes.

[230] Chi, above n 222, 22. Chay used 'peer recognition' as one of two criteria for selecting participants in his doctoral research on lawyer problem solving. The other criterion was years of experience. However, Chay did not use partnership and sign-off responsibility to indicate such recognition, but rather the fact that his two experts had been awarded specialist accreditation by their law society. Chay, above n 159, 81.

However, the choice of sign-off responsibility may not always be an appropriate point of distinction. Domain experts do not only exist amongst those professionals with such authority. It is plausible and to an extent inevitable that there are non-partners who perform equally as well as – and in some instances even better than – professionals who possess sign-off responsibility. There are, for instance, non-partners who may today be on the brink of promotion to partnership, save for the finalisation of promotion formalities. There are also non-partners who may be experts in a technical sense, but who lack the interpersonal skills needed to attract new clients to their organisation and expand their practices, and on this ground they fail to secure promotion.

It is also possible that some partners will not be experts in a specialist legal sense, but will have obtained sign-off responsibility on the strength of other personal attributes or professional abilities. This situation could occur either as a conscious act of the partnership promotion committee, or because of a phenomenon noted by Ericsson – with specific references to research by Ericsson and Lehmann on stock market advice, psychotherapy counselling and weather forecasting[231] – that 'people recognized by their peers as experts do not always display superior performance on domain-related tasks.'[232] In other words, peer confirmation may be expressly or impliedly based on traits other than subject-matter expertise – or be mistakenly bestowed on merely presumed domain experts.

Further, simply being a partner in a law firm or a senior manager in an economics firm or a competition authority, should not be a sufficient condition for scoring a participant positively under this measure. What is necessary is that their appointment to partnership or its equivalent level of authority be a reflection *at that time* of their specialist expertise in competition law. Only in this way can the argument be made that the relevant individual's specialist legal skills were attested to by their peers.

2 *10,000 Hours*

The second measure of likely expertise used to rank participants in this study was whether or not an individual had 10,000 hours of experience in competition law.

[231] Ericsson and Lehman, above n 226.
[232] Ericsson, above n 226, 4.

According to the so-called 10,000-hour rule as defined in a 1993 journal article by Ericsson, Krampe and Tesch-Romer[233] and subsequently popularized in Malcolm Gladwell's 2008 book, *Outliers*,[234] an individual who has engaged in 10,000 hours (or approximately 10 years) of deliberate practice in a specific area of endeavour can be considered an expert in that area. There have since been other scholarly articles written by a range of researchers, notably Kaufman[235] and Freeman,[236] who have relied on, defended or sought to confirm the validity and predictive power of this rule or a variation of it.[237] The origins of the rule itself can be traced back to the work of Watson in the 1920s, who contended that expertise – and even genius – can be developed in almost anyone[238] and is not dependent on an individual's innate ability as had been claimed by Galton in the 1860s.[239]

In addition to Gladwell's book, at least eight other bestsellers[240] have been written during the last decade expounding Ericsson et al's findings that 'high levels of deliberate practice are necessary to attain expert level of performance' and that 'dramatic differences in performance between experts and amateurs-novices' are in most cases (allowing for activities such as sports where the genetics of body size are important) due 'to similarly large differences in the recorded amounts of deliberate practice.'[241]

---

[233] K A Ericsson, R T Krampe and C Tesch-Romer, 'The Role of Deliberate Practice in the Acquisition of Expert Performance' (1993) 100 *Psychological Review* 363.
[234] Malcolm Gladwell, *Outliers: The Story of Success* (Little, Brown and Company, 2008).
[235] Kaufman welcomed Ericsson et al's work arguing that it provides 'a scientific way forward for research on giftedness, and offers exciting new ways to further our understanding of the determinants of high ability within a particular domain of expertise.' S B Kaufman, 'Investigating the Role of Domain General Mechanisms in the Acquisition of Domain Specific Expertise' (2007) 18 *High Ability Studies* 71, 71.
[236] J Freeman, 'If You Can't Measure It – It doesn't Exist' (2007) 18 *High Ability Studies* 65.
[237] According to Web of Science, Ericsson, Krampe and Tesch-Romer, above n 233, has been cited over 1,000 times.
[238] JB Watson, *Behaviorism* (The University of Chicago Press, 1930).
[239] F Galton, *Hereditary Genius* (Macmillan, 1869).
[240] Daniel J Levitin, *This is Your Brain on Music: The Science of a Human Obsession* (Plume/Penguin, 2007), Daniel Coyle, *The Talent Code: Greatness Isn't Born. It's Grown. Here's How* (Bantam, 2009), David Shenk, *The Genius in All of Us: Why Everything You've Been Told about Genetics, Talent and IQ is Wrong* (Image Books, 2010), Geoffrey Colvin, *Talent is Overrated: What Really Separates World-Class Performers from Everybody Else* (Portfolio Trade, 2010), Daniel H Pink, *Drive: The Surprising Truth About What Motivates Us* (Riverhead Books, 2011), Matthew Syed, *Bounce: Mozart, Federer, Picasso, Beckham and the Science of Success* (Harper Perennial, 2011), Daniel Kahneman, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011), David Brooks, *The Social Animal: The Hidden Sources of Love, Character and Achievement* (Random House, 2012).
[241] Ericsson, Krampe and Tesch-Romer, above n 233, 392.

Yet not all scholars agree with the significance of this rule. For example, Hambrick et al have recorded how 'Ericsson and colleagues' view has been roundly criticized on conceptual and methodological grounds,'[242] and that 'there is widespread skepticism … over Ericsson and colleagues' strong claims regarding the importance of deliberate practice for acquiring expert performance.'[243] Ackerman, one such critic, recently questioned the strong view of Ericsson et al's position stating that elite performance is not attainable by everyone, such that:

> some individuals may have accumulated 15,000h, 20,000h or more of deliberate practice, yet remain also-rans on the leader board, or never achieve medal status in national or international competitions.[244]

In their own research, Hambrick et al found that based on a review of six previous studies on the correlation between deliberate practice and chess-playing ability, 'on average, deliberate practice explained 34% of the variance in performance after correcting for measurement error … leaving 66% of the variance unexplained and potentially explainable by other factors.'[245] The authors go on to claim that not only do some people never become experts, regardless of how many hours or years they devote to practicing deliberately in a single area, there is clear evidence that 'people *do* reach an elite level of performance without copious practice.'[246] These findings accord with Gardner's contention that Ericsson et al's view on the sufficiency of deliberate practice 'requires blindness to ordinary experience.'[247]

Such findings and related practical issues raise questions about the applicability of the 10,000-hour rule as a definitive measure of specialist legal expertise. First, there is the difficulty of assessing the quality of work a competition law specialist has engaged in during their years of experience in the field. Deliberate practice is defined by Ericsson et al as 'engagement in highly structured activities that are created specifically to improve performance in a domain through immediate feedback, that require a high

[242] David Z Hambrick, Frederick L Oswald, Erik M Altmann, Elizabeth J Meinz, Fernand Gobet and Guillermo Campitelli, 'Deliberate Practice: Is That All It Takes To Become An Expert?' (2013) 45 *Intelligence* 34, 36.
[243] Ibid 37.
[244] Phillip J Ackerman, 'Nonsense, Common Sense and Science of Expert Performance: Talent and Individual Differences' (2013) 45 *Intelligence* 6, 9.
[245] Hambrick et al, above n 242, 38.
[246] Ibid 44 (emphasis in the original).
[247] H Gardner, 'Expert Performance: Its Structure and Acquisition' (1995) 50 *American Psychologist* 802, 802.

level of concentration, and that are not inherently enjoyable.'[248] To what extent a lawyer or economist dealing with competition law matters may have experienced all or even some of these things during their specialisation is difficult to determine, even if relying on personal responses to a detailed questionnaire.

There is also the possibility, as foreshadowed by Hambrick et al,[249] that some of the participants in this study with less than 10,000 hours of deliberate practice as a competition law specialist may nevertheless be experts under Hoffman's Scheme. To apply the 10,000-hour rule as a necessary precondition to being assessed as such would therefore seem arbitrary, and in some instances misleading. Nevertheless, it is not unreasonable to accept that this rule has some determinative significance, at least as one of five measures of likely expertise within the ranking-framework for this thesis.

The most acute issues were anticipated to arise when assessing those participants who had more than 8 years but less than 12 years specialist competition law experience. At the lower end of this range it was possible that each year qualified as deliberate practice according to Ericsson's formulation, and that this may have been sufficient for some individuals to rise to expert status in as few as, say, eight and a half years. At the higher end, a participant with eleven and a half years' experience may appear to have satisfied the 10,000-hour rule, yet might not have developed their cognitive skills as much as their less experienced colleagues. As Blasi noted with reference to the work of Avram Sherr,[250] it has been not uncommonly observed 'that some people have twenty years' experience while others seem to have but one year of experience twenty times.'[251]

3 *Conceptual Depth*

The third measure of likely expertise was the extent to which a participant demonstrated an ability to engage in deep conceptualisation while considering competition issues in a legal-risk assessment context.

---

[248] Hambrick et al, above n 242, 36.
[249] Ibid 44.
[250] Blasi referred to a conference paper presented by Sherr at a 1996 conference organized by the Australasian Professional Legal Education Council. Most of the central themes of that paper were repeated in Avrom Sherr, 'The Value of Legal Experience in Legal Competence' (2000) 7(2) *International Journal of the Legal Profession* 2.
[251] Gary L Blasi, 'Teaching/Lawyering as an Intellectual Project' (1996) 14 *Journal of Professional Legal Education* 65, 66.

Chi, Feltovich and Glaser's landmark study of novice and expert physicists demonstrated how greater expertise can be reflected in an increased focus on 'second-order, derived cues.'[252] In a more recent review of the relevant literature, Chi reiterates that experts can be distinguished from novices insofar as they 'perceive the "deep structure" of a problem or situation.'[253] This assessment comports with Lesgold et al's findings that experts are able to identify patterns, arrangements and cues that less expert practitioners do not perceive.[254] It is also consistent with the findings of Nievelstein et al who concluded in a study of law students and academics that their expert participants 'mentioned significantly more central concepts' when, like Chi et al's physics experts, they were given the think-aloud problem-solving task of grouping concept cards based on text-book problems.[255] Weinstein, who compared lawyers and law students in the area of social security disability law, similarly found that his expert subjects saw 'deeper patterns' while the novice participants 'tended to focus on surface features.'[256]

Guided by these studies, conceptual depth was chosen as the third measure for distinguishing between different levels of expertise amongst the participants in this study. More particularly, it seemed reasonable to expect that the more expert a participant in this study, the more likely they would be 'to use "abstract" representations that rely on "deep knowledge"; that is, imaginal and conceptual understanding of functional relations and physical principles that relate concepts.'[257] It was also assumed that they would be less likely to display the novice's tendency 'to use hastily formed "concrete" (that is, superficial) problem representations'[258] which lack the 'conceptually richer and more organized'[259] representations associated with greater expertise.

---

[252] M T H Chi, P Feltovitch and R Glaser, 'Categorization and Representation of Physics Problems by Experts and Novices' (1981) 5 *Cognitive Science* 121, 150.
[253] Chi, above n 222, 23.
[254] A Lesgold et al, 'Expertise in a Complex Skill: Diagnosing X-ray Pictures,' in Michelene T H Chi, Robert Glaser and Marshall J Farr (eds), *The Nature of Expertise* (Erlbaum, 1988) 311.
[255] Nievelstein et al, above n 219, 1055.
[256] Weinstein, above n 164, 6 footnote 16.
[257] Robert R Hoffman, 'How Can Expertise be Defined? Implications of Research from Cognitive Psychology' in Robin Williams, Wendy Faulkner and James Fleck (eds), *Exploring Expertise: Issues and Perspectives* (Macmillan Press Limited, 1998) 81, 88.
[258] Ibid.
[259] Ibid.

Two test cases in this study (Case A and Case B) provided the best opportunity to identify those participants who considered issues at a conceptually deep level and those whose assessments involved more superficial analyses. While neither of these cases was selected for this specific purpose, they were viewed as better suited to this task than the other cases because they permitted a clearer conceptualization of different levels of analytical abstraction. The other cases were problematic because for Case C there was an option for participants to bypass substantive analysis altogether using various inference strategies, and for Case D the range of agriculture-related products and services in that case resulted in a complex and potentially unwieldy factual matrix not easily framed in terms of conceptual depth.

Given that conceptual depth would be only one of five different measures used to distinguish between different levels of likely expertise in this study, the potential detriment of using just two rather than four test cases for this measure was reduced. Moreover, the ultimate aim was not so much to rank individual participants conclusively, but to identify, in broad terms, the degree to which they were likely to be experts. This invocation of a second-order analysis further reduced concerns about unobserved performance against this one measure.

As to the analysis of the relevant test cases themselves (which are more fully described in Part D below and in Attachment B), Case A concerned the supply of large precast concrete products used in road works and other infrastructure projects. Both the acquirer and the target company manufactured reinforced concrete box culverts and drainage pits, which were the key overlapping product-lines offered by both merger parties and were the focus of the ACCC's inquiries. These were the basic facts of the case. As with all the test cases, the market definition task then required participants to identify what they considered to be the relevant product (or service) and geographic dimensions of the relevant market.

In this particular test case, participants' analytical approaches to defining a relevant market were considered likely to be determined in large part by whether or not they viewed the relevant concrete products primarily in terms of their superficial characteristics (that is, the substantial weight of large, reinforced concrete pipes and arches) or their functionality (that is, the utility and use of the relevant products in various contexts). Where the weight of the concrete products proved to be the focus,

participants would be more likely to devote their time to assessing how far such heavy cargo might be transported. However, where functionality – the second-order feature of the relevant products – was their focus, participants would be more likely to spend their time analysing product substitutability.

Functional substitution is a key issue in merger cases, as alternative products or services need to be identified before geographic issues are considered. Without knowing which other types of products or services are likely to constrain the pricing and supply decisions concerning the products or services supplied by the merging parties, it is impossible to identify the locations and availability of those sources of constraint, and hence the geographic extent of the market.

Experienced practitioners typically seek first to understand as much as they can about other products or services that could perform the same functions as those supplied by the merger parties. Lay people and less expert competition law specialists, on the other hand, are more likely to rely on the obvious (but superficial) fact that concrete products are heavy with a low value-to-weight ratio. Accordingly, such individuals can be expected to devote more time and effort to analysing issues relating to the feasibility of transporting these products by various means and across various distances. They are less likely to perceive that lighter, corrugated steel substitute products, for instance, might be sourced from a wide variety of suppliers and be viably transported over far greater distances, thereby concluding that a widened geographic market would be more appropriate. Such an assessment could materially affect competition-law risk.

On this basis, those participants choosing to spend relatively more time – and rely more – on geographic considerations in their assessment of Case A, were viewed as adopting a more superficial analytical approach than those who emphasised product substitution analysis. Participants relying on arguments that there may be no geographic overlap between the merger parties – and hence no competition concern at all – would therefore be considered less likely to be experts on this measure than those who focused more on supply and production-substitution possibilities, which in conceptual terms were more abstract and more directly associated with greater analytical depth.

Case B concerned the merger of two freight-forwarding businesses. The acquiring business (but not the target business) operated as a vertically integrated freight-forwarding and shipping service provider though a joint venture arrangement it had

with a shipping company. Shipping services were central to this case because the main transportation routes at issue included a 240 kilometre sea-crossing between the Australian mainland and the island state of Tasmania.

It was anticipated that some participants would fail to detect a potentially problematic vertical-integration issue because they viewed both merger parties as logistics companies of roughly equivalent structure. Others might identify freight-forwarding services as the relevant service dimension, and on that basis view the merger as essentially a horizontal merger without a vertical dimension. Their analysis would therefore be 'deeper' than those who either did not identify the significance of this distinction between freight-forwarding and integrated logistics operators. However, this second group of participants would still not have undertaken the deepest level of possible conceptualisation.

Participants would engage in deeper conceptualisation if they recognised that freight-forwarding was the relevant focus for identifying the potential anti-competitive effects of the proposed merger, but that foreclosure to this market could arise from the acquirer being vertically integrated and potentially motivated to restrict the supply of its shipping services to third-party freight forwarders post-merger. These participants' ability to discern this so-called 'vertical risk' while other participants might either overlook or not fully grasp the significance of the asymmetrical economic structures of the merger parties, was considered an indicator of the former's greater expertise inasmuch as their analysis required a deeper conceptualisation of the facts of the case. Not only would these participants identify the relevant structural differences between the parties, they would also identify the behavioural implications of this distinction.

To summarise with reference again to Chi, Feltovich and Glaser's study of physics expertise, those participants who viewed the concrete products in Case A primarily in terms of their physical characteristics would be like the authors' novice physicists who grouped physics problems according to their descriptive characteristics, rather than according to underlying physics principles as the experts did. Whereas the less expert competition law specialists in this study would 'basically use the features explicitly stated in the problem'[260] and be more likely to identify problem distinctions 'on the

[260] Chi, Feltovich and Glaser, above n 158, 150.

spot,'[261] the more expert participants could be expected to focus more on issues of functionality. In doing so, they would be guided by the economic principles of product substitution while relying less on the more obvious physical characteristics of large reinforced, concrete products. Their focus – and hence their verbal protocols – would therefore dwell proportionally more on evidence of equivalent functionality in the form of steel products and products made from other materials.

Those participants considering Case B would have the option of relying on the literal descriptions of the parties in the ACCC's letter, which, as it happened, had poorly explained the differences between the merger parties (the letter read as if the acquirer and target companies serviced different customers, when in fact they did not). Alternatively, they could focus more on the nature of the commercial activities implied in the letter and develop a richer representation of the relevant services. Those participants who took this latter route would be more likely to avoid errors arising from the ACCC's mismatched comparisons and recognise vertical integration issues as central to a correct assessment. In doing so, they would demonstrate a deeper conceptualization of the case.

4 *Exceptional Reasoning*

The fourth measure of likely expertise was the extent to which a participant engaged in exceptional reasoning when considering competition issues in a legal-risk assessment context.

Given the range of specialists participating in this study, not only was it assumed that some participants would perform better, and others worse, but that certain individuals would perform very differently from all other participants, either in a positive or a negative sense. While the test cases themselves contained limited information thereby precluding a direct comparison of ultimate opinions since many participants may simply be unable to provide a conclusive view, unique reasoning strategies would be readily identifiable. Accordingly, those participants whose reasoning strategies were significantly and materially different from those of every other participant would be identified as having engaged in exceptional reasoning.

[261] Ibid.

This measure of likely expertise was foreshadowed by Weinstein, who expressly limited his terminology to 'experienced' and 'inexperienced' solvers.[262] He chose not to use the 'expert' designation because it connoted 'exceptional performance' amongst experienced legal specialists, whereas his focus was on differences attributable to participants being either a novice or an 'outstanding practitioner' without further dividing the latter category in terms of degree of expertise.[263] At the same time, however, he observed that exceptional answers arrived at by some experts, but not by others, could be a distinguishing feature of greater expertise and be worthy of future research.[264]

To determine whether or not these unique reasoning strategies could be considered positive or negative outliers, an ideal strategy or response was formulated for use as a conceptual template against which the quality of a reasoning strategy could be assessed. Those exceptional performances that conformed closely to the ideal would be ascribed a positive rating, while those that bore little resemblance would be rated negatively. An ideal response according to this measure of likely expertise was considered a unique reasoning strategy which:

i. The participant relied on to a material extent when assessing the case (trivial reasoning differences were therefore excluded);

ii. Was grounded in known facts, experience and inferential insight, rather than superficial distinctions and dubious assumptions;

iii. Reflected extensive specialist knowledge in competition law analysis; and

iv. Was novel, creative and compelling.

Comparisons against an ideal response are not new in studies like the present one. They have been used extensively in the wisdom studies conducted by the Berlin Group, amongst other researchers.[265] Some of the most recent endorsees of this approach

[262] Weinstein, above n 164, 24 footnote 114.
[263] Ibid.
[264] Ibid.
[265] The ideal response methodology used by the Berlin Group as detailed in Ursula M Staudinger, Jacqui Smith and Paul B Baltes, *Manual for the Assessment of Wisdom-Related Knowledge* (Max Planck

(which, like this study, relied directly on the Berlin Group's *Manual for the Assessment of Wisdom-Related Knowledge*[266]), were Greaves et al in their 2014 study of transformational leadership styles.[267] That study, and several of the earlier studies cited by the authors, were similar to the present study insofar as it combined measurements against an ideal response and think-aloud problem solving and verbal protocol analysis.

This study applied principles from the Berlin Group's Manual when formulating the above ideal response to assess reasoning quality in the competition law problem-solving tests used here. That document also informed the approach to scoring and weighting responses against this ideal, as discussed below in Section 6.

5 *Comprehension Errors*

The fifth and final measure of likely expertise was the extent to which a participant made comprehension errors when considering competition issues in a legal-risk assessment context.

Comprehension errors included misreading the test-case documentation or recalling irrelevant or inaccurate information in an attempt to augment that incomplete information. Those participants who made more non-trivial errors, especially where these were material or determinative in a problem-solving sense, would be assumed less likely to be experts. This was considered a non-controversial assumption supported by Chi's observation that it can be generally expected that experts will solve-problems with fewer errors than lesser experts within the expert's specialist knowledge

---

Institute for Human Development and Education, 1994, was used in J Smith, UM Staudinger and PB Baltes, 'Occupational Settings Facilitating Wisdom-Related Knowledge: The Sample Case of Clinical Psychologists' (1994) 62 *Journal of Consulting and Clinical Psychology* 989, PB Baltes, UM Staudinger, A Maercker and J Smith, 'People Nominated as Wise: A Comparative Study of Wisdom-Related Knowledge' (1995) 10(2) *Psychology and Aging* 155, M Pasupathi, UM Staudinger and PB Baltes, 'Seeds of Wisdom: Adolescent's Knowledge and Judgment About Difficult Life Problems' (2001) 37 *Developmental Psychology* 351. One of the most recent studies using this methodology is CE Greaves, H Zacher, B McKenna and D Rooney, 'Wisdom and Narcissism as Predictors of Transformational Leadership' (2014) 35(4) *Leadership & Organization Development Journal* 335.

[266] Ursula M Staudinger, Jacqui Smith and Paul B Baltes, *Manual for the Assessment of Wisdom-Related Knowledge* (Max Planck Institute for Human Development and Education, 1994).

[267] CE Greaves, H Zacher, B McKenna and D Rooney, 'Wisdom and Narcissism as Predictors of Transformational Leadership' (2014) 35(4) *Leadership & Organization Development Journal* 335.

domain.[268] In the present study, material errors were therefore assumed to be associated with lower levels of expertise.

The documentation used in this study was in a form that was unfamiliar to most participants. It was also purposely insufficient to support a fully-reasoned response to the legal-risk assessment task. Further, as a result of the procedural rules applied in this study, no participant would have prior familiarity with any of the test-case transactions.[269] Assessing the difficulty of securing clearance for a specific merger transaction based on these documents alone would therefore be a novel experience.

Given these factors, it was considered inevitable that errors of various kinds would occur, but that the more expert participants would be less prone to serious errors. As Colon-Navarro had found, the experts in his study were not misled by what he termed 'deliberate red herrings implanted in the hypothetical' backstory he had prepared for his interviewee.[270] This, he argued, attested to their ability to distinguish between relevant and irrelevant information, and to self-correct when possible ambiguities were discovered.

6 *Scoring and Weighting*

A scoring and weighting scheme was developed to facilitate the ranking of participants based on the above five measures of likely expertise. Ranking required the summation of scores across each of the five measures. These final scores would form the basis for rankings according to participants' levels of likely expertise.

The following sections (a) to (c) describe the methodological choices made to establish a theoretically coherent and consistent approach to quantifying the relative levels of likely expertise amongst those individuals volunteering to participate in this study. This quantification process focused first on participants' possession of sign-off responsibility and 10,000 hours of experience as legal specialists. It then shifted to the scoring and weighting of participants' performances according to their observed

[268] Michelene TH Chi, 'Laboratory Methods for Assessing Experts' and Novices' Knowledge,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 167, 170.
[269] The procedures by which test cases were disqualified if a participant stated that they had prior knowledge of the details of the case, is explained in Part D of this chapter.
[270] Colon-Navarro, above n 217, 131.

conceptualization ability, their exceptional reasoning ability, and their avoidance of comprehension errors in an information-constrained and time-limited context.

(a) Sign-off Responsibility and The 10,000 Hour Rule

The two experience-based measures of expertise (appointment to partnership and 10,000 hours of specialist experience) were viewed as essentially binary in nature. A participant either had the experience (or status) or they did not. The only complication, as noted previously, was when it could not be ascertained conclusively whether or not either scenario was satisfied. For instance, a participant may have been a partner with full sign-off responsibility for several years, yet may not have been appointed to the partnership on the basis of his or her specialist experience in competition law. Accordingly, a participant's partnership status could not be ignored, but their promotion to partnership also needed to be a reflection of their then competition law expertise. Partnership promotion per se could not be presumed to be a peer-based attestation of their expertise in a specialist knowledge domain unless this condition was satisfied.

Similarly, some participants may have had 10 years' experience as a competition law specialist, yet during their first few years (or in more recent times) they might not have devoted all their time to competition law matters. Alternatively, they might not have engaged in increasingly demanding competition law work that would enable them to grow their knowledge and abilities in a deliberate-practice sense. It was therefore considered preferable to avoid blindly equating their 10,000 hours of experience with the 20,000 hours of experience of another participant whose position at the most senior levels of their profession largely validated the quality of their experience in building expertise.

In addition, there existed no previous research confirming partnership and sign-off responsibility to be unambiguous indicators of legal expertise. The so-called 10,000 hour rule, as previously discussed, also remains controversial inasmuch as expertise may develop in substantially less time for some individuals, but take much longer (or may never develop) for others.

For these two measures of likely expertise, three possible grades were chosen. First, a +1 score depicted clear possession of the required experience. For partnership this

equated to promotion to partnership after having been a specialist competition lawyer or economist for more than eight years prior to that promotion. For the 10,000 hour rule, it equated to a participant having developed their expertise as a competition law specialist over a period of at least 12 years.

At the other end of the scale, not being a partner with sign-off responsibility was scored as an automatic -1, even when the participant in question could have been on the brink of partnership promotion (which was not something that could have been realistically determined during pre-test interviews[271]). Similarly, specialist experience of eight years or less scored -1 for the reason that this represented a substantial discount to the 10,000 hours rule, even assuming this rule to be overstated in some instances.

For those participants where neither a +1 nor a -1 could be confidently given, a '0' score was assigned to reflect an unobserved or inconclusive assessment. This was considered an appropriate compromise where a participant had been promoted to partnership but had had eight or fewer years of prior experience as a competition law specialist, or had only gained specialist status after becoming partner. A score of '0' was also given when a participant had more than 8 but fewer than 12 years' experience as a competition law specialist. This range was intended to accommodate borderline cases, with a not insubstantial margin of error on both the high and low sides of the conventional 10-years cut-off period. It also allowed for participants having only loosely approximated their responses to the pre-test questionnaire.

The choice to restrict these partnership and experience measures to a range from -1 to +1 reflected the view that the correlation between partnership promotion and level of expertise has not been conclusively established, and that based on available scholarly research the 10,000-hour rule remains controversial. It was therefore considered necessary to avoid placing excessive weighting on either of these measures, particularly given that other performance-based measures would also be used and an appropriate calibration across all measures could become over-complicated if a wider range of scores had been chosen.

[271] As it happened, at least one non-partner participant was promoted to partnership within 12 months of being tested in this study.

(b) Performance-Based Measures

The performance-based measures of conceptual depth, exceptional reasoning and comprehension errors, were not considered binary, but rather measures of degree. The deeper a participant conceptualized the core issues in a case, for example, the greater their apparent level of likely expertise. It was therefore decided to give greater range to this measure than was chosen for assessing partnership status or 10,000-hours of experience. It was also decided that greater weighting could be given to these tests of performance because they were blind to any preconceptions based on the participant's status or seniority, and therefore were not prejudiced by the assumptions associated with those experience-based measures.

The better an exceptional reasoning strategy was assessed to be, the higher (and weightier) a participant's likely expertise rating. Similarly, an exceptionally poor reasoning strategy required a correspondingly substantial lacking-in-likely-expertise rating. As to comprehension errors, it was important to distinguish between errors that were trivial and errors that were substantial (either in seriousness or number, or both). There was therefore a need to reflect these differences in more nuanced terms.

At the same time, it was necessary to maintain uniformity between all the measures used to rank participants. If partnership status and more than 12,000 hours' specialist experience were both to be scored at +1, then significant conceptual ability should not be scored +10, or at least not if the ultimate aim was to arrive at a compelling summation of likely expertise scores across all five measures. The imperative of intra-measure meaningfulness therefore needed to be balanced against the need for inter-measure consistency.

Moreover, while a -1 to +1 score (for partnership and '10,000 hours' experience) and a -5 to +5 score (for each of deep conceptualization, exceptionally good reasoning and no comprehension errors) may appear a significant improvement over the much wider -10 to +10 range, there would still exist a range of only three points in the former measures (-1, 0, +1) compared to a substantially more extended (and more difficult to justify in a proportional sense) range of eleven points across each of the latter measures (-5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5).

A conservative weighting of the performance-based measures from -2 to +2 was therefore considered appropriate insofar as it permitted 5 degrees of total range (-2, -1, 0, +1, +2) with two degrees of positive performance above the neutral or unobserved grade of '0,' and two degrees of negative performance below. This arithmetic combined with the conceptualization of each measure relating to levels of likely expertise rather than actual expertise, translated into the following scoring and weighting rules.

(i) Conceptual Depth

Conceptual depth was assessed by analysing the issues identified by participants as they considered two test cases, Case A and Case B. As previously noted, Case A concerned the merger of two manufacturers of reinforced concrete products. Depending on the degree to which participants focused on the functionality of the products in question, both in absolute terms and relative to their consideration of geographic issues, they would be scored as follows:

a) Participants who undertook a balanced analysis of (product) functionality and (geographic) characteristics were scored at '0.' This neutral score would also be given to participants who did not consider this case;

b) Those participants who placed significantly greater focus and relied more on geographic substitution (indicating superficial analysis) were scored at -1. If they went on to develop arguments to suggest that there might be no geographic overlap between the parties (thereby indicating that the participant had relied predominantly on an analysis of the *characteristics* of the products rather than their *functionality*), they were scored at -2; and

c) Participants who focused more on product functionality and supply-side substitution, were scored at +1. Those who relied predominantly on product and production substitution (with minimal consideration of, or reliance on, geographic issues), were scored at +2.

Case B concerned interstate freight-forwarding inclusive of a shipping component over the sea channel separating mainland Australia from Tasmania. Depending on the

degree to which participants were able to identify and separate the vertical components of the relevant services – and then whether or not they identified the risk of vertical foreclosure arising from the vertical integration of the acquirer – they would be scored as follows:

a) Participants who identified that the central issues in this case would revolve around increased market concentration in freight-forwarding services, leading them to characterize the transaction in purely horizontal merger terms, would be scored at '0.' The same neutral score would be given to participants who did not consider this case or whose views could not be clearly ascertained;

b) Participants who characterized the relevant services as integrated logistics or shipping services, were scored at -2. This score reflected the fact that such a characterization would have the effect of automatically precluding the consideration of the deeper vertical foreclosure risk; and

c) Participants who identified a potential vertical foreclosure issue whereby there was a risk that the integrated acquirer may restrict access to freight forwarding after the merger was completed, were scored at +2.

There was no clear method for assigning scores at -1 or +1 given the ostensibly limited nature of the conceptual analysis that could be undertaken by participants in this case. In this regard, Case B permitted a less fine-grained assessment of conceptual depth compared to Case A. However, it was still important that the deepest level of conceptual analysis in both cases be scored equally at +2. Hence, the only scores possible for Case B analyses were -2, '0' and +2.

(ii) Exceptional Reasoning

As noted in Section 4 above, an ideal response according to this measure would be a unique reasoning strategy which:

i. The participant relied upon in a material sense when assessing the case (trivial strategies were excluded);

ii. Was firmly based on known facts, experience and inferential insight, rather than superficial distinctions and dubious assumptions;

iii. Reflected extensive specialist knowledge in competition law analysis; and

iv. Was novel, creative and compelling.

The assessment of participants' performances against this ideal response was based on the Berlin Group's 7-point scale as set out in its 1994 Manual.[272] The first level of that scale (Level 1) denotes a response with 'very little' similarity with the ideal response. The mid-level rating (Level 4) denotes a response that is 'moderately' similar. The highest rating (Level 7) indicates responses that are 'a great deal' similar to the ideal response.

Importantly, instructions in the Manual identify the avoidance of extreme ratings as a natural approach for some raters, but one that should be avoided. Raters are advised to suppress such 'reservations towards extreme praise or extreme criticism' in unnecessary attempts 'to avoid false judgments.'[273] As the authors explain, the assessment process is not one of judgment in a subjective or values sense, but of performance against the objective criteria set out in the ideal response.[274]

In the present study, exceptional reasoning strategies that bore 'very little' similarity to the four elements of the ideal response would receive a -2 score. Strategies that were 'moderately' similar would be rated at '0' – or at -1 or +1 if a slightly less or slightly more positive case could be made, respectively. Strategies that were 'a great deal' similar to the ideal response would be rated at +2.

The decision to limit this study's scale to 5 points (-2, -1, 0, +1, +2) rather than the Manual's original 7-points (1, 2, 3, 4, 5, 6, 7) was done to ensure consistency with the other measures. There were no obvious disadvantages associated with this change, apart from a loss of potential (though in practice very little) interpretive richness compared to a 7-point scale.

[272] Staudinger, Smith and Baltes, *Manual for the Assessment of Wisdom-Related Knowledge*, above n 265, Appendix 2.
[273] Ibid 27.
[274] Ibid.

(iii) Comprehension Errors

Comprehension errors were assessed across all the cases considered by each participant. Not only were the seriousness and materiality of individual errors relevant, but also the number of times errors were made in a cumulative sense.

Participants who made two or more trivial or non-material errors would be scored at -1. One or more material comprehension errors, that is, errors that could be considered detrimental to a participant's overall assessment of a case, would result in a score of -2.

Participants whose only error related to misstated information or poorly expressed descriptions in the provided documentation (most notably the ACCC's misleading comparison of the merger parties in Case B) would be scored at '0,' even though the misunderstandings that resulted were in some instances material. Participants who overcame or corrected for such misstatements and bad descriptions (and who otherwise made no significant errors) would be scored at +1.

(c) Summary

The solicitation of participants for this study involved a managed sampling of both partners and non-partners in order to ensure a range of specialist legal expertise amongst volunteer participants. However, this simple binary distinction was considered insufficient for ranking participants according to their likely levels of expertise. The advantage of this distinction was that it was useful for volunteer-solicitation purposes and was broadly related to expertise levels. But using it to rank participants for the purposes of subsequent testing in the course of this study would have been clearly sub-optimal.

Having identified five measures of likely expertise – inclusive of a modified partner and non-partner rule – the relevant definitions and relevance of these measures could be presented in conceptual or abstract terms. However, their effectiveness depended on a workable, and theoretically defensible, scoring and weighting methodology. Absent a means of translating the relevant theoretical concepts into quantifiable measures, the ultimate aim of ranking participants according to their levels of likely expertise would not have been possible.

The above scoring and weighting procedures were conservatively, yet logically, framed. First, the uncertainties surrounding the significance of sign-off responsibility and the 10,000-hour rule were contained within a three-step scoring range of -1 to +1. This delimited the remaining three performance measures to a five-step scoring range of -2 to +2 (for measures three and four) and from -2 to +1 (for measure five), thereby facilitating intra-score flexibility while ensuring inter-score compatibility. The result was a maximum cumulative-score range of -8 to +7 across all measures, which was deemed broad enough to capture relative differences with sufficient nuance while avoiding excessive distance between participants with the highest and the lowest overall scores.

Secondly, the use of five measures reduced reliance on any one measure, while the concept of 'level of likely expertise' further ameliorated concerns regarding unobserved behaviours. At the same time, the above scoring and weighting system could be expected to produce ranking results that would be efficacious in the context of the study's overall objectives. In these regards, the above methodological choices were considered sufficiently innovative to overcome the novelty of this aspect of the study, and yet workable in terms of relevant theoretical and practical requirements.

D *Test Procedure*

The procedure for testing the cognitive skills of study participants was central to this study's overall design. It was amongst the first issues to be considered in a conceptual and theoretical design sense. But it would become the last step in the front-end process by which suitably qualified volunteers were solicited, ranked and tested. In terms of procedural sequence, therefore, it is appropriate to discuss the testing process now.

The issues discussed below relate to the instructions that guided participants through the testing process, the decision to use actual merger review case documentation (instead of hypothetical merger case materials), the significance of the information-limited and time-constrained design of the representative task, the manner in which the specific test cases were selected, and how various practical issues were addressed using largely technological solutions to enable the remote testing of busy professionals via the telephone and Internet.

1 *Task Instructions*

The representative task in this study required participants who specialised in Australian competition law to assess legal-risk based on market inquiry letters relating to four pending merger investigations which were at the time being actively considered by the ACCC.[275] These letters – which were addressed generically to interested parties – were downloaded from the mergers page of the ACCC's website and presented to participants according to procedures described further below. For the participant who specialised in New Zealand competition law, documents in the form of Statements of Preliminary Issues, which are roughly equivalent to the ACCC's market inquiry letters, were taken from the website of the CC.[276]

Each participant was provided written instructions as to their specific task. These instructions included a summary of the methodological basis for the testing process in which they were participating. This information was included in the website used to guide participants through their interviews (see further Sections 5 and 6 below). The relevant webpage on which these written instructions were provided is reproduced in Figure 3.1 below.

[275] The ACCC is able to review mergers under a number of different formal and informal procedures that provide various degrees of statutory or practical immunity from legal action under Section 50 of the *Australian Competition and Consumer Act* 2010. The market inquiry letters used in this study were issued as part of the ACCC's informal merger clearance process as described in its *Informal Merger Review Process Guidelines*. See further http://accc.gov.au/ .

[276] The Commerce Commission's Statements of Preliminary Issues were, in this study, more detailed pre-decision documents than the ACCC's market inquiries letters. The CC issues these documents pursuant to formal procedures under Section 66 of the New Zealand *Commerce Act* 1986. For the purposes of this study, the analytical task faced by this participant was for all practical purposes identical to that faced by all other participants, save that geographic and other localizing features were different. See further http://www.comcom.govt.nz/ .

FIGURE 3.1 – Methodology Page from Test Website

METHODOLOGY

'Think-aloud' analysis involves reading a document analytically and vocalizing your thoughts as they occur to you. In this study, you will engage in think-aloud analysis as you read through and assess preliminary merger clearance cases.

Your Task:

To assess (hypothetically) the level of difficulty involved in securing merger clearance. As you consider the issues relevant to your assessment for each case, talk constantly – as if to yourself.

Importantly, do not try to explain why you are thinking about certain issues, just that you are thinking about those issues. This will provide the best evidence of how you naturally think. Explanations require additional cognitive processing (such as formulating coherent sentences and ensuring the logical presentation of ideas). This could distract you from your task and compromise your usual analytical approach.

Why This Methodology?

Think-aloud analysis overcomes a key problem in studies of this kind, namely: Experts often do not know what or how they do what they do, and even if they do know, they cannot explain it. This is why traditional interviews and questionnaires alone have historically yielded very little information about how experts really think.

By thinking aloud without trying to make sense of your vocalizations, you will generate verbal data that can be studied using scientific methods (such as verbal protocol analysis) to develop a better understanding of the cognitive elements of applied competition law analysis.

Keep Talking – To Yourself

What you say during the following exercises may sound incoherent and even confused. This is normal and is in no way a sign of any lack of expertise. Analytical thinking often involves apparent leaps in logic, repetitive questioning, periods of doubt and even confusion. These are all attributes of expert analysis and precisely what we are seeking to observe.

Almost certainly, you will need to work against your training and experience as a professional adviser. When speaking to clients, we never want to sound inarticulate or unsure of ourselves. When speaking with junior staff, we have to avoid sounding confused or incoherent. But in this exercise, an uncensored, 'stream of consciousness' mindset is required.

Because it can feel unnatural to 'think aloud' while reading a case, you may sometimes stop talking. If this happens, you will be prompted to 'keep talking' so that your thoughts continue to be vocalized and recorded.

Apart from this prompting, you should proceed as if you are alone and talking to yourself.

The instruction in this Methodology Page for participants to 'assess the level of difficulty in securing merger clearance' was a formulation developed during the pilot

tests for this study.[277] The original instruction was simply to predict the outcome of a case, that is, whether or not it was likely to be cleared by the relevant competition authority. This is similar to the test stated in Figure 3.1 insofar as it also involves a predictive element. In practice, however, assessing the level of difficulty of a case was considered more realistic inasmuch as – according to comments made by one pilot study participant and other participants involved in the actual study – clients commonly ask for such advice when assessing the commercial and legal risks associated with their merger proposals.

Moreover, using test instructions that ask for a level of difficulty discourages binary thinking in which a 'pass' or 'fail' response might be offered. Assessing difficulty levels seemed to involve – and arguably required – a deeper analysis of risk elements leading to greater reflection on the indeterminacy of specific issues. This was considered likely to produce richer vocalizations.

2 *Use of Actual Merger Review Cases*

As noted in Part A of this chapter, in some competition law jurisdictions the competition authority has a role in reviewing, authorising and clearing merger proposals, either formally or informally. These authorities typically publish the basic facts of pending merger proposals for the purpose of inviting public comment on possible issues of concern. While this is not a universal practice,[278] it is the usual approach adopted in the jurisdictions in which the present study was conducted.[279]

These pre-decision publications typically offer a factual overview of the proposed transaction and identify potential competition issues. They may also include a list of questions about the scope of relevant markets and related considerations, and can

[277] Two experienced competition law specialists trialled different wordings/instructions on the study website. They also assisted in estimating how much time the representative task would take to complete and therefore how many test cases could be complete in one 45-minute interview.
[278] For instance, merger review practice in the UK, at the time of this study, permitted non-statutory, informal merger advice from competition officials, but the publication of the facts of a given case and the seeking of public comment prior to providing such advice were not features of this process.
[279] In addition to the public review process which is the focus of this study, the relevant competition authorities also undertake confidential reviews. Some of these merger proposals are subject to pre-assessment by authority staff, and may be quickly cleared without any public review on the basis that there is no material competition issue to consider further. This study is only concerned with those transactions that are in the public domain and which competition officials felt they could not clear summarily, assuming that they were approached for a pre-assessment decision first. See further the Australian Competition and Consumer Commission, *Informal Merger Review Process Guidelines* (Australian Competition and Consumer Commission, September 2013) and the New Zealand Commerce Commission, *Mergers and Acquisitions Guidelines* (Commerce Commission, July 2013).

sometimes state or imply the merger proponents' arguments as to why their transaction should be allowed to proceed. Within weeks – though sometimes months – of publishing such documents, competition officials announce their decision either to not oppose the transaction or to undertake a more in-depth second-stage review, which may ultimately jeopardise the merger proponents' plans. The majority of transactions subjected to informal review are permitted to proceed, with a small number having clearance conditions attached to them. Few mergers are opposed following a public review, and even fewer end up in litigation.[280]

As previously discussed, not only does this area of competition law practice draw on specialist legal skills, merger cases themselves are relatively easy to describe in summary form and do not require time-consuming analyses of a generic kind. The only remaining issue from a study design perspective was therefore whether to use hypothetical or actual merger review cases.

In this instance, the use of actual cases was preferred over hypothetical scenarios. First, actual cases, in the form of merger review letters and Statements of Preliminary Issues published by the ACCC and CC, respectively, were readily available in the two competition law jurisdictions familiar to the participants in this study. Second, the selected cases provided realism without known outcomes, which presented the opportunity to compare the accuracy – or usefulness – of participants' opinions. However, this would also be a double-edged sword. The impending nature of these cases meant that all participants would need to be tested within a four to five week time-frame or otherwise the published results of a case could unfairly benefit later participants and confound subsequent analysis.[281] This restricted window of opportunity would inevitably limit the number of study participants, although the

[280] For an overview of these elements of the ACCC's merger review processes see Australian Competition and Consumer Commission, *Informal Merger Review Process Guidelines* (Australian Competition and Consumer Commission, September 2013). For a guide to the New Zealand Commerce Commission's approach, see its *Mergers and Acquisitions Guidelines* (Commerce Commission, July 2013).

[281] For the four ACCC test cases used in this study, the number of calendar days from the date of a merger clearance application to the ACCC's final decision ranged from 57 to 72 days. The market inquiry letters used in this study were issued between three and four weeks after the merger clearance applications had been lodged. This limited window of opportunity was not an issue for the New Zealand expert who was the sole participant from that jurisdiction and therefore the test cases he assessed only needed to remain undecided up to the day of his testing.

logistics of testing up to 20 participants were considered to be a potentially more significant limiting factor.[282]

Hypothetical cases, on the other hand, would have needed to be prepared by the researcher, whose editorial preferences would have been unavoidably reflected in the final documentation, even if authored by a third party. In addition, hypothetical scenarios can suffer from artificiality, insufficient complexity (or over-complexity) and superficial factual depth. There is also the risk of unforeseen inconsistencies and critical gaps existing within a fictional factual-matrix. Further, distracting ambiguities and participants obsessing over minor technical errors are further consequences of mishandled hypothetical case studies. In combination, these negative potentialities could have substantially compromised the validity and richness of the resulting verbal data.

3 *Information-Limited and Time-Constrained*

The experiment-like nature of the representative task as explained to participants was integral to its design. Participants were purposely given limited information about each case and were constrained by the time allotted for them to provide their opinions. The extensive use of limited information tests by experimental psychologists in real-world settings has been documented by Hoffman,[283] Schweickert et al,[284] and Shadbolt and Burton.[285] According to Crandall, Klein and Hoffman, tests of this kind are particularly useful 'in probing the specialized sub-domain knowledge or reasoning of experts.'[286]

The risk in the present context was that such a task could prove too artificial and all that would be tested was participants' strategizing about an unrealistic task that could not be completed because of unrealistic information and time limitations. According to

[282] Based on previous studies of this broad type (which have often involved teams of researchers working together), a general expectation was that a minimum of fifteen participants in this study would strike a reasonable balance between generating useful data sets and ensuring feasibility for a single researcher working alone.
[283] RR Hoffman, 'The Problem of Extracting the Knowledge of Experts from the Perspective of Experimental Psychology' (1987) 8 *AI Magazine* 53.
[284] R Schweickert, AM Burton, NK Taylor, EN Corlett, NR Shadbolt and AP Hedgecock, 'Comparing Knowledge Elicitation Techniques: A Case Study' (1987) 1 *Artificial Intelligence Review* 245.
[285] NR Shadbolt and AM Burton, 'Kowledge Elicitation Elicitation Techniques: Some Experimental Results,' in KL McGraw and CR Westphal, *Readings in Knowledge Acquisition* (Ellis Horwood,1990).
[286] Crandall, Klein and Hoffman, above n 140, 104.

Newell,[287] Jenkins[288] and Crandall, Klein and Hoffman,[289] this is a recognised methodological issue, not just for think-aloud verbal protocol analysis but for psychological testing generally. Hence, the objective became to create a task that was controlled in information availability and in timeframe, but which was capable of yielding 'data that possess ecological relevance, validity, and representativeness.'[290]

The information contained in the ACCC's market inquiry letters and the CC's Statements of Preliminary Issues was not overly detailed, yet it was considered sufficient to provide enough background for participants to work with. These documents consisted of up to half a dozen pages of text which, as was confirmed in the pilot tests, could be scanned and considered in sufficient depth within a loosely-imposed 10-minute time limit. Any fewer pages would likely have meant materially insufficient information. More than half a dozen pages, on the other hand, could have resulted in information overload in the sense that participants could have legitimately spent all their allotted time simply reading through the documentation. The Statements of Preliminary Issues published by the CC came close to this limit. However, these documents also included large portions of standardised wording which could be – and were – largely ignored or glossed-over by participants.

A 10-minute time limit per test-case was stated but not actively enforced so to assess the self-monitoring skills of participants and specifically their ability to sense elapsed time and finish within the time limit – a skill that legal professionals may develop under a billable hour system where every 6 or 10-minute block is counted as a billable unit. It also added a degree of pressure to the legal-risk assessment task, which was considered likely to add to the realism of the task in the sense that lawyers in practice often have limited time to form their preliminary opinions. This limit had the added advantage of facilitating an expectation amongst participants that they would be able to complete up to four test-case assessments during their 45-minute to one-hour interviews.

[287] A Newell, 'You Can't Play a Game of 20 Questions with Nature and Win' in WG Chace (Ed), *Visual Information Processing* (Academic Press, 1973), 283.
[288] JJ Jenkins, 'Four Points to Remember: A Tetrahedral Model of Memory Experiments' in L Cermak and F Craik (eds), *Levels of Processing in Human Memory* (Lawrence Erlbaum & Associates, 1978) 429.
[289] Crandall, Klein and Hoffman, above n 140, 91.
[290] Ibid.

4 *Test-Case Selection*

At the beginning of the testing phase of this study, 10 ACCC market inquiry letters and four Statements of Preliminary Issues by the CC were selected as possible test cases. These cases were, at the time of participant interviews, listed in chronological order (according to submission date) on the respective websites of the two competition authorities as merger transactions currently under review. Accordingly, they were neither hypothetical nor selectively chosen in the sense that the researcher's preferences might otherwise have been reflected in determining which transactions the participants would ultimately consider. Moreover, they were not unusual cases and nor were they concerned with obscure transactions in a substantive sense when compared with the completed review cases also listed at the time on the competition authorities' websites.[291]

These 10 market inquiry letters and four Statements of Preliminary Issues were uploaded to the study website where they could be viewed by participants during the testing process. At the researcher's direction, participants clicked a link on a relevant page of the website which made the first letter or Statement of Preliminary Issues appear on their computer screens. Participants were then asked to confirm whether or not they had any direct or indirect knowledge of the transaction described in the document. This question was easily answered once a participant had briefly scanned the first page of the letter or statement. Any such knowledge disqualified that test case, leading the participant to close the opened webpage and select another letter until they found a transaction about which they had no special knowledge. General knowledge concerning the identity of the parties or familiarity with the industry concerned was not a disqualifying condition.[292]

[291] The industries involved in these test cases were, in a broad sense, similar if not the same as those industries in which well-known court cases had been decided, such as those involving building products, transportation services, and agricultural and livestock services. In other words, these cases were more a 'staple' of competition law jurisprudence than aberrant or atypical.

[292] Each participant was also required before-hand to complete a practice exercise and a practice case prior to undertaking their think-aloud analysis of these test cases. The practice exercise involved a participant thinking aloud while counting the windows in their home. This exercise was based on the practice exercises described in the Appendix to Protocol Analysis. The practice case was a market inquiry letter or Statement of Preliminary Issues relating to a merger transaction that had already been cleared by the ACCC or CC, respectively. The purpose of this preparation, which followed the procedures detailed in the Appendix to Protocol Analysis, was to ensure participants' familiarity with the think-aloud problem solving technique. It also permitted the researcher to explain further certain aspects

As anticipated in the design of the case selection process, the first participant in this study – who was first by virtue of being the first to volunteer and arrange a time to participate in the study[293] – disqualified the first two ACCC market inquiry letters, but successfully completed his think-aloud analysis of the following four letters. These four transactions were then labelled Case A,[294] Case B,[295] Case C[296] and Case D.[297] Subsequent Australian-law participants were instructed to analyse only those four cases, which they all did with the exception of two participants not completing Case A, five not completing Case B, two not completing Case C and one not completing Case D.[298] The New Zealand competition law specialist did not consider any of these cases, but rather analysed three of the four CC cases originally selected from its website. These transactions were described in the Statements of Preliminary Issues labelled cases E, F and G.[299]

Summaries of the facts and outcomes of each test case are provided in Appendix B.

5 *Remote Interviews*

The remote interviewing of participants – in this instance via long-distance telephone calls – was an arrangement the researcher considered was supported both theoretically and empirically, although no similarly distant arrangements were found in the

---

of the methodology, including reminding participants of the importance of their continuous vocalization while assessing legal risk for each case.

[293] This participant was the first individual to volunteer for the study, and in this sense his self-selection was a random event or at least one that did not reflect a choice on the part of the researcher.

[294] Case A – Rocla Pty Ltd – proposed acquisition of Beresford Concrete Products Pty Ltd (Informal Review 51157, Date Commenced 23 April 2013, Date Completed 19 June 2013). http://transition.accc.gov.au/content/index.phtml/itemId/1110327/fromItemId/750991

[295] Case B – Toll Holdings Limited – proposed acquisition of Linfox Trans-Bass business (Informal Review 50982, Date Commenced 22 March 2013, Date Completed 30 May 2013). http://transition.accc.gov.au/content/index.phtml/itemId/1115976/fromItemId/751046

[296] Case C – Brambles Industries Limited – completed acquisition of CEVA Limited's Pallecon business (Informal Review 50545, Date Commenced 19 March 2013, Date Completed 24 May 2013). http://transition.accc.gov.au/content/index.phtml/itemId/1115051/fromItemId/751046

[297] Case D – Ruralco Holdings Limited – proposed acquisition of Elders Rural Services Limited (Informal Review 50798, Date Commenced 19 March 2013, Date Completed 30 May 2013). http://transition.accc.gov.au/content/index.phtml/itemId/1115891/fromItemId/751046

[298] The primary reason for a participant not considering a case was their prior knowledge of the transaction. A secondary reason was lack of time for those participants who took longer completing the pre-test questionnaire or the practice case.

[299] Case E – Perry Metal Protection Limited/ CSP Coating Systems, 24 December 2012, withdrawn 24 May 2013 – http://www.comcom.govt.nz/clearances-register/detail/781; Case F – Bertelsman SE & Co. KGaA and Pearson plc, File Number 14007, 20 December 2012, Cleared 19 March 2013, Decision Number [2013] NZCC 6 – http://www.comcom.govt.nz/clearances-register/detail/780; Case G – Bligh Finance Limited and Hire Equipment Group Limited (Hirepool/Hirequip), File Number 13913, 16 October 2012, Cleared 21 February 2013, Decision Number [2013] NZCC 2 – http://www.comcom.govt.nz/clearances-register/detail/772.

literature. In Protocol Analysis the authors note that a researcher need not be present during the think-aloud problem solving process, although historically this was necessary so that he or she could monitor participants' performances and prompt them to 'keep talking' when needed.[300] Even so, Ericsson and Simon posited that a timed buzzer or similar device might be adequate for this purpose.[301]

With advances in technology there is no longer a compelling reason for the personal presence of the researcher – and potential benefits in them being absent as confirmed by Klinger's research in the 1970s.[302] Christensen, in a more recent think-aloud study testing legal reading skills, recorded how she intentionally left her test subjects alone because she believed it would increase the reliability of her data.[303] Her approach was consistent with Ericsson and Simon's view that social interaction is not helpful during the performance of the representative task because in such a setting 'demands for coherence and reasonable completeness' can compromise the otherwise 'close correspondence between verbal protocol and the actual processes used to perform the task.'[304]

In the present study, the advantages of being able to test a larger number of highly-specialist, time-constrained professionals distributed across Australia's major commercial centres outweighed any concerns over the apparent novelty of using telephone interviews to guide participants through a website containing standardised instructions and test materials. The fact that the researcher was able to prompt participants effectively whenever they stopped speaking (which was rarely required), further indicated that the fundamental requirements of an effective think-aloud problem solving test were met. Moreover, no participant indicated that this approach was uncomfortable for them, presumably because they were all experienced advisers

[300] Ericsson and Simon, *Protocol Analysis*, above n 139, 83.
[301] Ibid.
[302] E Klinger, 'Utterances to Evaluate Steps and Control Attention Distinguish Operant from Respondent Thought While Thinking Out Loud' (1974) 4 *Bulletin of the Psychonomic Society* 44.
[303] 'Although many think alouds are done with the researcher in the room, I specifically left the room because I felt that both the experts and law students (who were sensitive to being judged or evaluated) would read more naturally if they performed the think aloud alone. As such, I believe me being absent for their think aloud increased the reliability of the data.' Leah M Christensen, 'The Paradox of Legal Expertise: A Study of Experts and Novices Reading the Law' (2008) *Brigham Young University Education and Law Journal* 53, footnote 56.
[304] Ericsson and Simon, *Protocol Analysis*, above n 139, xv.

familiar with discussing complex matters via the telephone, albeit in this instance they were required to speak to themselves.[305]

All participants were adroit at scrolling up and down the on-screen documents and seemed generally comfortable with this aspect of the test process. Some participants vocalized their decisions to jump back to earlier pages, while it was apparent from a review of the transcripts that others were jumping back and forth quite freely albeit without a situational commentary. While it is likely that most participants would have preferred hard copy documents if given the choice, the approach adopted did not of itself create obvious problems in terms of document navigation, comprehension or reasoning processes. Indeed, it may well have resulted in greater vocalization of cognitive activity.

6 *Technology*

The recording and analysis of participants' think-aloud vocalizations relied heavily on several technologies, some of which were Internet-based while others took the form of specialised software. The primary challenge was the remote location of the researcher who was for the duration of the study based in Hong Kong, several thousand kilometres away from the study participants located in various cities in Australia. Other challenges related mainly to ensuring that the study could be conducted cost-effectively by a single researcher.

As previously noted, all interviews were conducted via telephone. This had a number of advantages. First it meant that testing could be completed remotely so no face-to-face meetings were needed. Logistics constraints were therefore limited to finding appropriate interview times that suited both the participant and the researcher. Secondly, as participants were typically in their own office at the time of their interviews, disruption to their day was minimised. Thirdly, last minute rescheduling

[305] It is noted, however, that during the pilot tests one participant indicated a preference for a hard copy document on which he could note his thoughts, draw diagrams and flip back and forth between pages. He explained that for him, drawing diagrams of ownership and market structures was part of his usual approach when considering these kinds of cases. Nevertheless, he subsequently confirmed that he was able to adapt to an on-screen review with little trouble, noting that this change probably led him to talk and reflect more on what he was doing. No other participants made similar comments – but this may have been only because their assigned roles did not include critiquing the test design. In any event, there was no evidence that this was an issue of concern for any participant.

was easily accommodated simply by changing the time of an interview.[306] Had in-person interviews been necessary, this would have potentially required re-scheduling the researcher's travel (and accommodation) arrangements and re-booking meeting rooms.

Because lawyers and professional economists have for decades conducted much of their work over the telephone, this was assumed to be a familiar medium, albeit there was the obvious novelty of having to think-aloud while engaging in legal-risk assessment tasks.

All telephone calls were initiated by the researcher using Skype-Out[307] mostly to participants' office landlines, at minimal cost.[308] Those participants who chose to be interviewed via Skype were interviewed without the video option to ensure consistency with those participants who only used the telephone. There were four times during interviews where the call connection failed. Fortunately, these temporary disconnections occurred during the preliminary stages of the calls and not during participants' analysis of test cases.[309]

To guide participants through the test procedures, a dedicated website was used in parallel with the telephone interview.[310] This websites was built for free on Wordpress.com,[311] although the time spent on constructing over 20 interconnected-webpages (each structured as an independent website to prevent participants independently 'surfing' through the test procedures) was not insignificant. This free on-line service permitted the creation of password protected pages which the researcher used to regulate the pacing of interviews and to confirm participants' consent to the conditions of their involvement, including their agreement to the audio recording of

[306] One law firm partner provided only 40 minutes' notice of his interest in being interviewed. This 'take it or leave it' offer was not missed because of the flexibility of being just a telephone call away and of having a live website (and associated audio recording software) available on-call 24 hours a day.
[307] http://www.skype.com/en/
[308] Total cost of Skype-Out calls made to participants in this study was approximately HKD120.00 or USD15.50.
[309] To minimize the disruptive effects of disconnection, a protocol was established at the commencement of each interview whereby participants were told that the call may be disconnected temporarily at any time. They were pre-advised to hang up their phones and wait for the researcher to call back. This did not prove to be a material problem in any instance. However, this possibility should be noted by any future researchers seeking to utilise Internet-based voice calls.
[310] http://taps28info.wordpress.com/
[311] http://wordpress.com/

their think-aloud problem solving.[312] All the test cases used in the study were uploaded to this website so that participants could open and close them as instructed.

One of the most important features of the website was that it enabled the use of standardised descriptions of methodology and procedures (such as the Methodology Page reproduced in Figure 3.1 above). This ensured that all participants were given the same information in the same sequence during their interviews. The introductory e-mails sent to prospective participants contained links to the first three pages of the study website which individuals could read at their convenience to learn more about the study and determine whether or not they wished to volunteer.[313]

The use of Skype enabled the telephone interviews with participants to be recorded using Call Graph software.[314] This free software started recording automatically once a call began and stopped when the call was ended. These recordings were formatted as MP3 audio files and stored on the researcher's password-protected computer hard drive. These files were subsequently transcribed using f4 transcription software,[315] which produced a text document that was linked to the original audio file by way of embedded time-code stamps. This meant that by clicking on any phrase in the text the associated audio would replay from that point.

The analysis software used was MAXQDA,[316] which converted the f4 documents into files that could be analysed with reference to observed performance and behavioural differences. By maintaining the pre-established connection between text and audio file, the researcher could review at any time how a phrase was vocalized, including the level of confidence and the general tone of a participant's response on specific issues. This was in addition to having available the full-text transcript of each participant's think-aloud verbalizations.

[312] To gain access to the website questionnaire and problem-solving tests, participants were required to confirm their consent to participate by clicking a 'Yes' button on the website Consent Form and then enter a password provided by the researcher.
[313] In addition to choosing an appropriately professional-looking 'skin' for the website, the official logo for The University of Queensland was displayed, with permission, on each page of the site. In accordance with ethical requirements, contact details of the University's Ethics Officer were also provided, which possibly increased the comfort levels of prospective participants.
[314] http://scribie.com/free-skype-recorder
[315] http://www.audiotranskription.de/english/f4.htm
[316] http://www.maxqda.com/

E *Analysis*

The post-testing phase of this study – the so-called back-end phase – required the analysis of the think-aloud verbal transcripts generated by study participants as they assessed the level of legal risk involved in securing clearance in each of the merger-review test cases. The first task in this phase involved grouping participants according to their levels of likely expertise. The second involved conceptualising a method for identifying apparent differences between how participants from each of these expertise groups tackled the chosen representative task. The third involved finding theoretical references from previous studies on which to construct coherent explanations as to the likely cognitive bases for the observed differences.

This part is divided into three sections. Section 1 briefly describes the manner in which participants were allocated to one of three groups according to their levels of likely expertise as determined from the scoring process described in Part C above. This description includes a discussion of a proposed alignment between participants' likely expertise scores (and group allocation) and the categorisation of expertise development outlined in Hoffman's Scheme.

Section 2 explains the choice of an exploratory approach to finding relevant differences between the assessment behaviours of participants from different groups. Because there were no previous studies with the same focus on legal specialists and analytical objectives, it was decided that a more open approach to categorising different behaviours would better serve the overall goal of identifying expertise-based cognitive differences.

Section 3 lists the findings of previous studies potentially relevant to developing explanations, based on cognitive analysis, for the results of this study. This includes the 3-step problem solving framework used to analyse the operations of participants long-term, short-term, and working memories, as well as a complementary analysis of participants' reliance on intuition and analytical reasoning. Reference is also made to previous studies which were considered relevant if not ultimately instructive in an analytical or interpretive sense.

1 *Participant Groups*

To facilitate the identification of cognitive differences associated with different levels of domain expertise, prior to testing each participant was ranked according to their cumulative scores of likely expertise as determined via the process discussed in Part C above. They were then allocated to one of three groups: Group A for the highest-ranked participants and Group C for the lowest-ranked participants. Group B participants would be those ranked in the middle ranks, between these two outer groups.

The choice of weighting parameters used to rank participants potentially yielded a maximum score of +7 and a minimum score of -8, for a total spread of 16 points. The grouping of participants according to whether they fell into group A, B or C – each of which contained a range of five score points – would be determined along the lines described in Table 3.1 below. This table also indicates how these groupings were assumed to correspond to the categories of apprentice, journeyman, expert and master under Hoffman's Scheme.

TABLE 3.1 – Participant Groupings According to Hoffman's Scheme

	APPRENTICES/JOURNEYMEN						APPRENTICES/JOURNEYMEN EXPERTS/MASTERS					EXPERTS/MASTERS				
RANKING SCORES	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7
	GROUP C						GROUP B					GROUP A				

As shown, Group A participants on the right-hand side of the table would be assumed to only contain masters and experts, Group C to the left only journeymen and apprentices, and Group B in the middle a mixture of all categories. Some participants in this middle group might perform more like experts or masters, while others could perform more like apprentices and journeymen.

Significantly, the interval between the lowest-ranked members of Group A and the highest-ranked members of Group C was six points (-3 to +3). This interval is greater than the result of any single measure of likely level of expertise used in the ranking process. This meant that the closest-ranked members from each of these two groups

were distinguishable across more than one measure. For members at the centre-points of these two groups (-5 and +5) the difference in likely levels of expertise was even more significant; and for those participants at the extremes the difference were greater still. The larger the intervals, the greater the level of confidence as to observed differences associated with different levels of likely expertise.

Within each group, however, it was more difficult to apply a ranking system. The separation of participants by just one or two points could not be viewed as a significant difference given the assumptions underlying the scoring methodology. Accordingly, it was decided that beyond placing participants within one of the above three groups, no intra-group ranking would be applied. This meant that all Group C participants, for instance, would be assumed (at least initially) to possess the same level of likely expertise. It also meant that an ordinal ranking of all participants based on their scores alone would be neither feasible nor necessary within the conceptual framework of the study.

2 *Identifying Differences*

This study was driven by a qualitative approach to identifying relevant behavioural differences between participants. This meant first examining any observed patterns of behaviour that appeared to be associated more with Group C or Group A participants. This approach contrasted with the prescriptions for pre-selected protocol codings based on 'entities (processes) postulated by a theory' and recorded ahead of time in a coding manual.[317] The ultimate focus of that methodological approach is protocol typology and the measurement of relative frequencies of procedural functions,[318] which is different from – and to a large extent inconsistent with – the objectives of this study.

That said, statistically significant differences between Group C participants and Group A participants would also be of interest in addition to purely qualitative differences. If populations within one of these groups displayed performance or behavioural characteristics that were statistically distinguishable from populations in the other group, this could be of greater significance than less common, idiosyncratic differences potentially attributable more to personality differences than differences associated with levels of expertise. Given the aim of securing the involvement of up to 20 volunteers

[317] Ericsson and Simon, *Protocol Analysis*, above n 139, 205.
[318] Ibid 204-205.

for this study, it was considered that statistically significant data could prove both meaningful and useful.

The analytical approach adopted for this study was by design exploratory. The aim was to identify – for the first time – cognitive differences between individuals with differing levels of expertise within a single specialist domain of law rather than to test any preconceived hypotheses about what form these differences might take. This led to the adoption of a more map-like approach – as opposed to an architectural, model-refining approach – to use Baddeley's characterisation of his own approach.[319] It also conformed to the investigatory principles espoused by Toulmin.[320] Viewed in this broader context, the objective of the present research was not to rely on the falsification of a prior theoretical position, but rather to contribute to 'generating fruitful questions that will increase our knowledge' as is arguably more productive in the early stages of theory development.[321]

Parallels with this approach can be found in the study of macrocognition, a term first associated with the work of Cacciabue and Hollnagel,[322] and subsequently Klein et al.[323] Klein et al state that macrocognition is 'a level of description of the cognitive functions that are performed in natural (versus artificial laboratory) decision-making settings.'[324] It is distinguishable from the study of microcognition which focuses on 'the building blocks of cognition,' which are general and invariant cognitive processes

[319] A Baddeley, 'Working Memory: Theories, Models, and Controversies' (2012) 63 *Annual Review of Psychology* 1, 3. Baddeley states, '… any complete theory is likely to require explorers in its initial stages and architects to turn broad concepts into detailed models. I myself am very much at the explorer end of the continuum, but I fully accept the importance of the skills of the architect if theory is to develop.'

[320] Toulmin, S, *The Philosophy of Science* (Hutchison, 1953). Toulmin's views were opposed by Popper who, according to Baddeley, 'argued strongly that a valid theory should make clear, testable predictions, allowing rival theories to confront each other in the all-important "crucial experiment" that settles the issue,' but who allegedly abandoned this view – along with Lakatos – in favour of views more sympathetic to Toulmin's map-like approach to theory development. Baddely, above n 319, 3-4. See also Popper, K, *The Logic of Scientific Discovery* (Hutchison, 1959). Baddeley suggests that Popper, too, 'subsequently abandoned the reliance on falsification.' Baddely, above n 319, 4.

[321] Baddely, above n 319, 4. Baddeley specifically attributed the philosophical basis of this approach to Lakatos, whose arguments in his book *Proofs and Reputations* were taken to be supportive of 'a map-like view of theory' along the lines advocated by Baddeley in his own work as a pioneering researcher and developer of multicomponent approach to working memory. See I Lakatos, *Proofs and Reputations* (Cambridge University Press, 1976).

[322] P C Cacciabue and E Hollnagel, 'Simulation of Cognition: Applications,' in JM Hoc, PC Cacciabue and E Hollnagel, *Expertise and Technology: Cognition and Human-Computer Cooperation* (Lawrence Erlbaum Associates, 1995) 55.

[323] D E Klein, H A Klein and G Klein, 'Macrognition: Linking Cognitive Psychology and Cognitive Ergonomics,' in *Proceedings of the 5th International Conference on Human Interactions with Complex Systems* (University of Illinois, 2000) 173.

[324] Gary Klein et al, 'Macrognition' (2003) 3 *Human-Centered Computing* 81, 81.

used in all types of decision-making and problem-solving.[325] According to the same authors, the 'naturalistic perspective' is an effective approach to studying macrognition insofar as it can enable the observation of subjects in real-world settings where data can be overwhelming (or too few), complexities unavoidable, goals ill-defined, some variables simply unknowable, and where 'research participants are domain practitioners rather than college students.'[326]

Klein, writing on his own, uses the term 'naturalistic investigation' to describe how researchers can use empirical research methods to identify phenomena and ideas at the early stages of inquiry, and how these phenomena can be tested in laboratory settings later.[327] He describes having undertaken his own investigations of this kind when lacking the ability to test hypotheses owing to not knowing what he might find – even though he knew enough to persist with the expectation that new phenomena, insights and ideas would be the likely result of his efforts.[328]

In a similar way, without the benefit of earlier research comparing the reasoning strategies of experienced legal specialists (especially where these individuals are selected and ranked according to a series of measures rather than by mere assumption), it was decided to adopt an exploratory approach approximating the perspectives described by the above researchers. The descriptive account of participants behaviour, which as argued by Lipshitz et al is an appropriate focus for research on naturalistic decision making,[329] was also favoured in this study as a way of avoiding what Elstein has described as prescriptions based on theories that preclude inferences from the actually observed behaviours of experts.[330]

At the same time, however, the think-aloud problem solving dimension of the present study distinguishes it from a purely naturalistic investigation favoured by Klein whose most recent treatise on the nature of insight relied on 120 case studies drawn from incidents recorded in the media and in 'books, especially those describing innovations

[325] Ibid.
[326] Ibid.
[327] Gary Klein, *Seeing What Others Don't: The Remarkable Ways We Gain Insights* (Public Affairs, 2013) 26.
[328] Ibid 27.
[329] Raanan Lipshitz, Gary Klein, Judith Orasanu and Eduardo Salas, 'Tacking Stock of Naturalistic Decision Making' (2001) 14(5) *Journal of Behavioural Decision Making* 331.
[330] Arthur S Elstein, 'Naturalistic Decision Making and Clinical Judgment' (2001) 14(5) *Journal of Behavioral Decision Making* 363, 364.

and discoveries.'[331] While he also reviewed interview transcripts, these were selected from investigations undertaken up to thirty years earlier and were not of a think-aloud problem solving kind – nor were they any more structured than retrospective accounts given by interviewees in response to probing questions.[332] In contrast, the present study implemented a structured information elicitation methodology; one that followed closely the techniques and guidelines developed by Ericsson and Simon in Protocol Analysis.

Lastly, this study was designed with a data-driven orientation. While the theoretical foundations described in this chapter were constructed in a manner consistent with Ericsson and Simon's prescriptions for effective think-aloud verbal protocol analysis, these authors also expressly acknowledged that new phenomena may be missed when a theory-driven orientation is followed too narrowly.[333] This is notwithstanding there can also be risks associated with 'the entanglement of data with theory,'[334] which is another reason why in a study such as this, more information rather than less should be recorded to facilitate the subsequent identification of potentially confounding or undermining factors. Where such factors are apparent to the researcher, they can then be addressed directly. If not, they can at least be subjected to scrutiny by others who can consider both methods and results together.

This commitment to describing and recording the entire research process followed Lundgren-Laine and Salantera's recommendation that this be done 'so that the reader is able to follow the solutions of the researcher.'[335] It also enables the researcher to explore salient issues and ignore less relevant ones, in the knowledge that such decisions could be subjected to detailed review.[336] In those instances where this study

[331] Klein, above n 327, 29.
[332] Ibid.
[333] Ericsson and Simon, *Protocol Analysis*, above n 139, 275. The authors state: 'Theory should not preclude the scientist from searching for new phenomena, or from paying serious attention to phenomena he hits on adventitiously. Many important scientific discoveries have been made in this way, with little prior guidance or discipline from theory.'
[334] Ibid 276.
[335] Helja Lundgren-Laine and Sanna Salantra, 'Think-Aloud Technique and Protocol Analysis in Clinical Decision-Making Research' (2010) 20(4) *Qualitative Health Research* 565, 572. For an expanded explanation of this approach in the field of medical research, see K Malterud, 'Qualitative Research: Standards, Challenges, and Guidelines' (2001) *Lancet* 483.
[336] On the justification of a selective approach to elicited verbal information, Lundgren-Laine and Salantra explain, 'When the objective of the study is to perform a task and solve problems related to the decision making of participants, it is possible to ignore events that the researcher considers irrelevant and not related to the particular performance under observation.' Lundgren-Laine and Salantra, above n 335, 570-1. While the authors listened to all the verbal data elicited from their test subjects, they ultimately

diverges from the prescriptions and guidelines for analysing verbal protocols by Ericsson and Simon, it was considered sufficient that the reasons for adopting the chosen approach were made explicit.[337] This was specifically the case in the adoption of Yang's contextual approach to protocol analysis over Ericsson and Simon's more traditional 'information processing style' of analysis.[338]

Yang's conception of 'a functional model and multi-layered categorization scheme, developed inductively from the transcripts themselves,'[339] was central to this study's methodological approach. It also provided a theoretical basis for the selective choice of protocols and the relaxation of assumptions relating to the need for coding schemes to be self-contained, involve mutually exclusive codings, and be comprehensive. This was necessary to facilitate practicability and to avoid the common criticism of coding schemes that are 'too reductive and mechanical in complex environments.'[340]

An ancillary advantage of this approach, which maximised the adoption of a more transparent, long-hand approach to protocol analysis described in the following chapters, was that should the theories and hypotheses developed during the course of this research ultimately prove invalid or misguided, at least the evidence relied upon will have been recorded with the possibility of being reused (or at least subjected to detailed review) by future scholars. This was another feature of this study that satisfied the need for contextual relevance and comported with Baddeley's empirical approach.[341]

'omitted those parts of the data that were not related to the focus of our study questions.' Lundgren-Laine and Salantra, above n 335, 571. See further, M W van Someren, Y F Barnard and J A C Sandberg, *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes* (Academic Press, 1994).

[337] Methodological variations of this kind, particularly as they related to the assumptions underlying Ericsson and Simon's analysis of verbal protocols, were treated in a manner consistent with Yang's perspective that 'the coding of protocols is an interpretive act which cannot be made in isolation without taking its situational contexts into consideration … [and] researchers [should] make great efforts to recognize, examine, negotiate and adopt the subjects' perspectives which leads to a clearer understanding of the subjects' complex mental processes.' Shu Ching Yang, 'Reconceptualizing Think-Aloud Methodology: Refining the Encoding and Categorizing Techniques via Contextualized Perspectives' (2003) 19 *Computers in Human Behavior* 95, 108.

[338] Ibid 109.

[339] Ibid, 101.

[340] Lundgren-Laine and Salantra, above n 335, 568.

[341] Baddeley, above n 319, 3. Baddeley explained that as a result of researchers abandoning specific areas of theoretical study, in developing his own theories he had resolved early in his career to base those theories 'very closely on the evidence, which would survive even if the theory proved totally wrong.'

3 *Cognitive Analysis*

The cognitive analysis in this study began with identifying performance and behavioural differences between participants as they sought to complete their assigned legal-risk assessment tasks. Where these differences (which in the first instance would appear as largely qualitative) could be aligned with differences in levels of likely expertise, those differences would be recorded. This was not intended to be a comprehensive survey given the limitations of the test design and given that this would be the first time that some of these differences had been identified. Rather, attention would be given to the most apparent and most clearly definable differences – and to those differences that seemed to have greatest distinguishing power between participant groups.

In terms of the granulatory of the analysis, it was decided to adopt a broad macrocognitive perspective rather than a microcognitive one, to use the terminology of Cacciabiue and Hollnagel.[342] Attention would therefore be given not so much to 'the building blocks of cognition,' but rather to 'the development of descriptive models of processes such as decision making, sensemaking and problem detection.'[343] This analysis would not be undertaken within a purely naturalistic setting, however, but through combining naturalistic elements (such as current merger review cases and actual review documentation) with laboratory-like efficiency and control (principally carefully conceived and administered think-aloud problem solving procedures).

To facilitate this exploratory approach, an investigatory methodology was chosen that would enable all aspects of the problem-solving process to be captured within a flexible yet theoretically sound framework. That framework, which is discussed in section (a) below, describes the three steps of effective problem-solving from the input of data from the environment, to the retrieval of relevant information from a participant's long term memory, to drawing inferences and forming opinions relevant to the task at hand. It was anticipated that using this approach would yield insights into the areas or stages of the problem-solving process where differences in legal expertise might manifest in observable forms.

[342] Cacciabue and Hollnagel, above n 322.
[343] Klein et al, above n 324, 82.

Having analysed the results from the application of the above framework to the transcripts of participants, it subsequently became apparent that the explanations for why these results were observed could be expanded. While the research question for this thesis had by this stage been largely answered, it was considered both relevant and meaningful to investigate the deeper cognitive drivers behind these differences. To this end, an analysis of participants' reliance on intuition and analytical reasoning as noted in section (b) below was also undertaken.

Lastly, the findings and results of previous studies were used as references for observed expertise traits, both to help identify what the underlying cognitive differences might be in a given instance and also to determine to what extent this thesis confirms the findings of earlier research. At the same time, it was not expected that any exact matches would be observed because of the broad approach adopted for this study, which was not designed to test any specific cognitive trait. Further background to this aspect of the analysis process undertaken in this study is provided in section (c) below, which discusses potentially relevant findings concerning general expertise traits (subsection (i)) and from expertise studies that have investigated specialist legal thinking skills (subsection (ii)).

(a) Three-Stage Problem Solving

Feltovich, Prietula and Ericsson conceptualise three distinct stages of effective problem-solving.[344] These stages, which are listed below, were considered an appropriate and sufficiently flexible analytical framework for the analysis of participants' cognitive strategies while engaging in this study's representative task. In addition, reference was made to Baddeley's findings concerning the role of working memory ('WM'), which acts as the 'central executive' for processing information present in short-term memory ('STM') as well as knowledge stored in long-term memory ('LTM').[345] These three stages of the problem-solving process – adapted for present purposes to the legal-risk assessment task presented to study participants – can be stated as follows:

[344] Paul J Feltovich, Michael J Prietula and K Anders Ericsson, 'Studies of Expertise from Psychological Perspectives,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 41, 58.
[345] See A Baddeley, 'Short-Term and Working Memory,' in E Tulving and F Craik (eds), *The Oxford Handbook of Memory* (Oxford University Press, 2000) 77; A Baddely, 'Is Working Memory Still Working?' (2002) 7(2) *European Psychologist* 85; Baddeley, above n 319.

Stage 1: Seeking and perceiving data from the environment, which in this study was restricted to the information provided in the market inquiries letters and Statements of Preliminary Issues documents that participants were required to read and assess;

Stage 2: Retrieving relevant information from LTM and combining it with environmental data within WM. This information included both technical legal knowledge (assumed to be common to all participants), knowledge of similar and analogous cases, and knowledge of relevant industries; and

Stage 3: Drawing inferences from the data provided and knowledge retrieved from LTM, both to form opinions on levels of legal risk and to determine the kinds or forms of further data required to complete a legal-risk assessment.

Unlike a written assignment or a formal opinion provided after a thorough consideration of all relevant issues, the verbal protocols of participants in this study were expected to reveal information relatable to each of the above stages. It would therefore be possible to identify when a participant was reading or clarifying the information provided in the test-case documentation (Stage 1), when they were identifying relevant issues based on their recall of facts and analogies from their LTM (Stage 2), and when they were synthesising from the provided data and drawing inferences about the level of difficulty the merger parties were likely to face in securing clearance from the relevant competition authority (Stage 3).

With the plan of assigning each participant to one of Group A, Group B or Group C (and therefore their approximate status as an apprentice, journeyman, expert or master based on their likely expertise scores), it would then be possible to compare their performances in each of these three areas with reference to their determined levels of expertise. Where participants in one group performed consistently differently from participants in another group in a particular area, further investigations could be undertaken into those differences with reference to the relevant stage of their analysis in which those differences were most apparent.

(b) Intuition and Analytical Reasoning

The emphasis on generating data using the above problem-solving framework yielded findings that both satisfied the study's research question and raised further questions that were not originally contemplated during the study design process. This possibility was anticipated by the study's exploratory approach. Of course, not all further questions could be accommodated within the confines of this thesis. It was therefore decided to focus only on investigating the more compelling indicators relating to differences in how participants with different levels of apparent expertise used their intuition and analytical reasoning.

For this purpose, the work of researchers involved in the cognitive analysis of decision-making and judgment was referenced. The central contribution in this context was the work of Kahneman and Tversky on the interaction of System 1 and System 2 thinking,[346] which distinction has been the subject of a large number of studies by other researchers[347] and was most recently popularised in Kahneman's 2011 book *Thinking, Fast and Slow*.[348] A more detailed introduction to this framework is provided in Part D of Chapter 6, which in chronological terms is appropriate given that this methodology was applied after the main findings of the thesis, as set out in the earlier parts of Chapter 6, were confirmed and found to be suggestive of further thinking differences.

(c) Expertise Traits

Studies by previous researchers who have used think-aloud verbal protocol analysis to study expertise and expert performance informed both the methodological design of this study and the decision to focus on certain aspects of the performances of participants. However, the relevance of these earlier studies was limited given the aim

[346] See A Tversky and D Kahneman, 'Judgment Under Uncertainty: Heuristics and Biases' (1974) 185 *Science* 1124; A Tversky and D Kahneman, 'Extensional vs Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment' (1983) 90 *Psychological Review* 293; D Kahneman and A Tversky, 'On the Study of Statistical Intuition' in D Kahneman, P Slovic and A Tversky (eds), *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press, 1982) 493.
[347] See as examples S Epstein, 'Integration of the Cognitive and Psychodynamic Unconcious' (1994) 49 *American Psychologist* 709; K R Hammond, *Judgment Under Stress* (Oxford University Press, 2000); L L Jacoby, 'A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory (1991) 30 *Journal of Memory and Language* 513; S Chaiken, S and Y Trope (eds), *Dual-Process Theories in Social Psychology* (Guilford Press, 1999); R M Hogarth, *Educating Intuition* (University of Chicago Press, 2001); D G Myers, *Intuition: Its Powers and Perils* (Yale University Press, 2002).
[348] Daniel Kahneman, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011).

of this study to compare experienced legal specialists (rather than novices) in the same field of legal practice.

As noted in Chapter 2, previous researchers have either examined cognitive differences outside the area of law or investigated how lawyers think but in general law-related terms, that is, outside a specialist area of law. Furthermore, studies of lawyers that have used the methodologies discussed here have invariably compared novices in a knowledge domain with domain experts, which is different to the present approach which seeks to analyse cognitive differences between specialists in the same knowledge sub-domain where even the most inexperienced participants could have been considered experts in most generic settings.

In addition, the small number of participants in many of these studies increased the likelihood that observed differences merely reflected the individualistic attributes of certain participants, rather than any group-wide phenomena. This issue becomes more problematic when novices are excluded from testing, as in this study. Whereas the differences between novices and experts can be both readily identified and directly associated with differences in domain knowledge, such distinctions are not as clear-cut when domain knowledge is no longer a variable. When the number of participants increases, however, larger sample-size comparisons become possible and tests for statistical significance become feasible. The results from this type of study are therefore less likely merely to reflect individual idiosyncrasies.

Given that the generalizability of findings was a key objective of this study, previous research which focused on the verbal protocols of just one or two individuals was considered less likely to be relevant. Nevertheless, there were two areas of possible – although potentially only indirect – relevance concerning possible cognitive differences between the participants in this study, namely, previously identified general expertise traits and specialist legal expertise traits.

(i) General Expertise Traits

The first area of potential relevance was previous research undertaken in the broader field of expertise and expert performance. Amongst the many scholars in this area, several have focused attention on specific expert traits that appeared of interest in the context of distinguishing different levels of legal specialists.

These researchers have shown how experts can display a deeper conceptual understanding of problems than novices do,[349] though sometimes they gloss over the details.[350] Experts can also demonstrate a better understanding of their own limitations[351] (and display more conservative tendencies when giving their opinions[352]). But they can also be over-confident in their own abilities[353] as well as in those of junior colleagues and novices considering the same issues.[354] Experts' ability to self-monitor may include automatically correcting comprehension and reasoning errors, while their conclusions and strategies are likely to be better than novices and intermediates,[355] although this may not always be the case.[356]

Experts have also been found to be able to see patterns in data that novices cannot;[357] and sometimes it is the atypical case – with rare or unusual facts or issues – that most effectively reveals these exceptional abilities.[358] At the same time, domain experts can be less flexible in their methods[359] and more dependent on contextual information.[360]

---

[349] See, for example, Chi, Feltovich and Glaser, above n 158.
[350] See, for example: J F Voss, G Vesonder and H Spilich, 'Test Generation and Recall by High-Knowledge and Low-Knowledge Individuals' (1980) 19 *Journal of Verbal Learning and Verbal Behavior* 651; H G Schmidt and P A Boshuizen, 'On Acquiring Expertise in Medicine' (1993) 5 *Educational Psychology Review* 205.
[351] See, for example: M T H Chi, R Glaser and E Rees, 'Expertise in Problem Solving,' in R Sternberg (ed) *Advances in the Psychology of Human Intelligence* (Vol. 1) (Erlbaum, 1982) 7.
[352] R R Hoffman, G Trafton and P Roebber, *Minding the Weather: How Expert Forecasters Think* (MIT Press, 2005).
[353] M T H Chi, 'Knowledge Structure and Memory Development,' in R Siegler (ed) *Children's Thinking: What Develops?* (Erlbaum, 1978) 73.
[354] P J Hinds, 'The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Prediction of Novice Performance' (1999) 5 *Journal of Experimental Psychology: Applied* 205
[355] For instance, the best chess players have been shown to generate the best moves, and in this respect perform noticeably better than intermediate players. See A De Groot, *Thought and Choice in Chess* (Mouton, 1965). See also G A Klein, 'A Recognition Primed Decision (RPD) Model of Rapid Decision Making,' in G A Klein, J Orasanu, R Calderwood and CE Zsambok (eds), *Decision-Making in Action: Models and Methods* (Ablex, 1993), 138. There can also be greater diversity of views amongst experts. This has been observed with expert auditors. See Bedard, J, 'Expertise and Its Relation to Audit Decision Quality' (1991) 8 *Contemporary Accounting Research* 198.
[356] As previously noted, experts do not always give the best answers and may perform worse than less expert individuals, even when operating within their areas of expertise. See Ericsson, above n 162, 4; Ericsson and Lehman, above n 226; Johnson, above n 226; R M Dawes, 'A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making' (1971) 26 *American Psychologist* 180.
[357] Lesgold et al, above n 254, 311.
[358] Geoffrey R Norman, Donald Rosenthal, Lee R Brooks, Scott W Allen and Linda J Muzzin, 'The Development of Expertise in Dermatology' (1989) 125 *Archives of Dermatology* 1063.
[359] R J Sternberg and P A Frensch, 'On Being an Expert: A Cost Benefit Analysis,' in RR Hoffman (ed) *The Psychology of Expertise: Cognitive Research and Empirical AI* (Springer Verlag, 1992) 191.
[360] Some research suggests that medical experts may make greater use of – and be more dependent on – contextual information, such as the medical history and social activities of patients, in arriving at their diagnoses. See P J Feltovich and H S Barrows, 'Issues of Generality in Medical Problem Solving,' in H G Schmidt and M L de Volder (eds) *Tutorials in Problem-Based Learning* (Van Gorcum, 1984) 128.

Experts also typically use less cognitive effort to solve problems within their field of expertise, often because they utilise more efficient and effective strategies[361] and sometimes because they can recall 'instant' solutions, thereby avoiding intermediate steps of analysis.[362]

(ii) Specialist Legal Expertise Traits

Studies that have directly investigated legal and law-related thinking skills, fall into two groups. The first consists of those investigations that have considered generic legal thinking skills, such as reading court cases, accessing information from reference databases and passing bar exams. As previously discussed, most of these studies have involved comparisons between novices (often law students) and presumed experts. While these studies are helpful in a broad sense, they shed minimal light on what to expect when a range of competent legal specialists are presented with the same problem-solving tasks grounded in their common knowledge-domain. Even so, it was considered possible that such studies could prove useful in a confirmatory sense insofar as these observed expert-novice differences might be evident amongst the participants in this study.

A second group of think-aloud verbal protocol studies involving lawyers have focused on specialist fields of law. As has been noted, these studies have investigated differences between participants with either a lot of or very little (and at times no) specialist domain knowledge. In this sense, these have been tests of knowledge rather than cognitive ability per se, since the least knowledgeable participants were, by design, limited in their ability to compete directly against the domain experts in the set problem-solving tasks which invariably required extensive domain-specific knowledge.

Colon-Navarro's study described how his four expert legal specialists were faster at diagnosing relevant legal issues, had more effective mental schema to deal with problems in their area, and possessed extensive procedural knowledge.[363] He also noted how they were able to develop better strategies and plans as to next steps for their client. Moreover, he made special mention of how one of his experts was able to draw

[361] See, for example, K J Gilhooly et al, 'Biomedical Knowledge in Diagnostic Thinking: The Case of Electrocardiagram (ECG) Interpretation' (1997) 9 *European Journal of Cognitive Psychology* 199.
[362] See, for example, W Schneider, 'Training High Performance Skills: Fallacies and Guidelines' (1985) 27(3) *Human Factors* 285.
[363] Colon-Navarro, above n 217.

on his knowledge of recent policy changes at a political level to aid assessment of the likely success of different options. However, two of the experts in Colon-Navarro's study had less than five years of specialist experience,[364] whereas in the present study, all participants had, as a minimum, five years specialist experience, with potentially more than half having no fewer than ten years. Further, there were to be no domain novices in this study.

Weinstein's three experts were simply described as 'outstanding practitioners' in SSD claims.[365] He compared the problem-solving abilities of these lawyers with those of a small group of novices and one experienced lawyer who had no specialist expertise in the relevant field of law. His findings were therefore potentially of limited reference value. Further, Weinstein's observations regarding this non-specialist but experienced practitioner lacking 'both substantive knowledge and domain-specific models for dealing with the much more ill-structure problem' in that study,[366] added little to the contributions of the other studies cited in this chapter.

As with Colon-Navarro's research, Weinstein's study provided methodological guidance for this study, particularly insofar as his approach to testing lawyers depended on appropriately designed representative tasks for the purpose of think-aloud problem solving. More substantively, he noted how his 'experienced solvers' were less likely to follow the format of the file documents, but rather organised the 'basic information according to a deeper, idiosyncratic pattern, building a schematic' based on a legal rules.[367] This contrasted with his 'inexperienced solvers' who were more likely to follow the order of evidence as set out in the file.[368] While all individuals invited to participate in the present study would likely have been classified as experienced solvers by Weinstein based on their years of specialist experience in competition law (apprentices and above), it was considered possible that these observed differences may persist beyond the apprenticeship stage.

Weinstein also discussed how his participants were distinguishable by the degree to which they were able to engage in forward (typically, expert) rather than backwards

[364] Ibid121.
[365] Weinstein, above n 164, 18.
[366] Ibid 52.
[367] Ibid 26.
[368] Ibid.

(novice) reasoning[369] (which is a commonly cited but not conclusive distinction in the expertise literature[370]), how his experts tended to set more realistic and efficient goals based on the information available,[371] and how inexperience was revealed in the imprecise and generalised use and recall of information.[372] Weinstein's analysis of the subexpert's performance was considered less relevant insofar as it focused on how this participant's reasoning strategies were affected by a lack of domain-specific knowledge of SSD claims.

F *Conclusion*

The methodology described in this chapter sought to avoid 'the temptation to study differences in performance between experts and novices because there are readily available tasks to measure such differences.'[373] It aimed instead to compare the performances of different levels of legal specialists, all of whom were equally knowledgeable – in terms of substantive law and relevant procedures – within the specialist domain of competition law. This would require both the voluntary co-operation of a sufficient number of suitably qualified competition law specialists, and the identification of a representative task that could test cognitive abilities at the higher end of the expertise spectrum (it was neither necessary nor appropriate to attempt to accommodate the lack of knowledge of any novice participants).

This methodology was developed to answer the research question guiding this study which is focused on finding readily identifiable and measurable cognitive differences associated with different levels of specialist legal expertise. As described, the chosen approach was comprehensive, extending from the identification of relevant theories of cognition and expert-knowledge solicitation, to technological considerations associated with the logistics of securing the co-operation of a sufficient number of suitably

---

[369] Ibid.
[370] See Chi, above n 222, 24; J H Larkin, J McDermott, D P Simon and H A Simon, 'Models of Competence in Solving Physics Problems' (1980) 4 *Cognitive Science* 317; V L Patel and D R Kaufman, 'Clinical Reasoning and Biomedical Knowledge: Implications for Teaching,' in J Higgs and M Jones (eds) *Clinical Reasoning in the Health Professions* (Butterworth-Heinemann Ltd, 1995) 117; But also see P Lemaire and R S Siegler, 'Four Aspects of Strategic Change: Contributions to Children's Learning of Multiplication' (1995) 124 *Journal of Experimental Psychology* 83, where the authors describe how experts may also engage in backwards reasoning where this is likely to be a more effective strategy in a given context.
[371] Weinstein, above n 164, 31.
[372] Ibid 38.
[373] Ericsson, above n 148, 231.

qualified volunteers capable of performing the representative task, which was designed specifically to test the essence of their specialist legal expertise.

The involvement of qualified participants was recognised at the outset to be a not insignificant challenge. The intrusive nature of the testing process and the lack of any direct benefits to otherwise busy professionals may explain in large part the historical lack of similar studies. Researchers who have adopted broadly equivalent methodologies have commonly resorted to using law students, who are traditionally a plentiful and willing source of test subjects for this kind of empirical research. The present study, however, set a number of minimum requirements such that all the individuals to be tested would be considered experts in most if not all previous studies that sought to test generic legal thinking skills.

This study utilised publicly available information and targeted e-mails – as well as the recommendations of other participants – to identify and secure the involvement of a range of participants. The use of web-based technologies facilitated long-distance interviews and enabled a broader geographical spread of participants compared to the less flexible logistics and higher implementation costs of in-person interviews. The flexibility of being able to source participants from across the major commercial centres in Australia facilitated the possibility of a viable sample size from both methodological and feasibility perspectives.

The representative tasks in this study were naturalistic inasmuch as they required participants to assess live merger review cases of a kind commonly handled by specialist competition law practices in major Australian and New Zealand law and competition economics firms. By providing incomplete information under time-constrained test conditions it was considered possible to use the set tasks to focus on the essence of expertise in this domain, namely, the exercise of applied economic analysis within a specific regulatory context. Strict adherence to the theoretical principles underlying think-aloud problem solving and verbal protocol analysis would permit the rigorous assessment and comparison of the cognitive processes revealed through the vocalizations of participants.

Merger review cases were identified as an appropriate context in which to devise a representative task to test participant's cognitive performances. This was based on an assumption that merger review matters are an area of activity that only competition law

specialists are capable of assessing from a legal risk perspective. The assessment activity itself would require both substantive analysis (including defining relevant markets and assessing market contestability) as well as procedural considerations relating to the administrative role of competition authorities, whose anticipated decisions would be the ultimate focus. The instruction to perform a legal risk assessment on the information provided reflected the real-world expectations of merger proponents seeking the advice of competition law specialists.

The test cases used in this study were selected via a process within which the researcher exerted minimal editorial or subjective influence. It began with downloading information about then current merger-review cases from the public registers on the websites of the ACCC and CC. The relevant market inquiry letters and Statements of Preliminary Issues were subsequently offered as possible cases from which participants could choose depending on their prior knowledge of the transactions in question. No participant knew or could have known the outcome of any case, and for the most part the industries and issues in each case were unfamiliar to every participant. At the same time, the cases themselves were neither obscure nor so technically demanding that any participant would be unable to proceed with their assessment for lack of technical or general commercial knowledge.

The method for ranking participants according to their levels of likely expertise utilised five different measures. Two of these measures were experience-based, namely, attested sign-off authority – or partnership promotion in law firm terms – and years of specialist experience. The other three were performance-based and focused on participants' conceptual abilities, instances of exceptional reasoning and examples of comprehension errors. These latter three measures required a preliminary assessment of participants' analysis of the test cases used in the study.

With reference to assumptions about the levels of likely expertise amongst members of the sample group, it was theorised that participants could be divided into three groups (groups A, B and C) according to their levels of likely expertise, and particularly whether they were likely to be experts/masters (Group A) or apprentices/journeymen (Group C), based on Hoffman's Scheme of traditional expertise development. Then, by identifying those analytical strategies suggestive of patterns of cognitive performance

more likely to be associated with only one of these groups, it was considered possible to relate such differences to participants according to their levels of likely expertise.

These identified differences would then be assessed within an analytical framework based on (a) the three stages of effective problem-solving described by Feltovich, Prietula and Ericsson and augmented by Baddely's research on memory functions, (b) the concepts of intuitive and analytical reasoning developed by Kahneman amongst other researchers in the field of decision-making and judgment, and (c) previous studies which have identified (i) general expertise traits (that is, relevant traits identified in studies not involving specialist lawyers), and (ii) specialist lawyer traits, chiefly those identified in studies by Colon-Navarro and Weinstein.

The methodological principles underlying the overall approach of this study were those most closely resembling the naturalistic perspectives of researchers such as Klein et al, with exploratory imperatives aligned with Baddeley's empiricism, Toulmin's philosophy of scientific investigation, and Klein et al's methods of observation and continual revision of conceptions rather than adherence to 'a single, predefined procedure.'[374] This was considered appropriate given that there had not previously been a study like this which had sought to investigate cognitive differences between higher-level legal specialists in the same field of law.

At the same time, the knowledge elicitation techniques used in this study followed the practices recommended by Ericsson and Simon in Protocol Analysis. In this respect, there was a strong conventionality in the methodology used and in the connections this study would have to numerous previous studies that have successfully generated data capable of supporting compelling theories of cognition in the investigation of expertise and expert performance.

Lastly, with the possible replication of this study in mind, it was noted that the actual ACCC letters used in this study are no longer available on the ACCC's website. On the one hand this creates an obvious problem regarding the availability of these documents for future researchers. On the other hand, it ensures that the novelty of these cases is preserved for future studies. Of the four ACCC merger review cases used in this study,

[374] Klein et al, above n 324, 83.

only the ACCC's decision in the last case was publicised by way of a media release.[375] The outcomes of the other cases were unreported. It is therefore highly probable that within a few months of their completion there will be little or no retained knowledge of them amongst a large proportion of potential future participants with the requisite levels of experience and specialisation in competition law analysis. The same test cases could therefore be used again, which could significantly enhance the comparability of results of any future research and those of the present study.

With this possibility in mind, a separate website has been created on which all of the documents used in this study can be accessed and downloaded by future researchers.[376] In addition, the website used for conducting the interviews has been preserved as an on-line reference for future researchers wishing to undertake similar studies – or who want to attempt to replicate the results of this one.[377]

The methodology described in this chapter forms the basis for discussion throughout the rest of this thesis. This discussion continues with the next chapter which records the profiles of the competition law specialists who volunteered to participate in the study. This includes an analysis of the differences and similarities between these individual legal specialists. Some of these differences were found to be statistically significant, and others not. The core of the chapter, however, is dedicated to applying the five measures of likely expertise developed in Part C of the present chapter, and to the calculation of resulting expertise scores and relative rankings. These rankings are then used to allocate individual participants to one of Group A, Group B or Group C as discussed in Part E above, which was necessary for the identification of performance and behavioural differences in chapters 5 and 6.

[375] Case D – Ruralco: http://www.accc.gov.au/media-release/accc-to-not-oppose-ruralcos-proposed-acquisition-of-elders-rural-services
[376] http://taps28nztest.wordpress.com/
[377] http://taps28info.wordpress.com/

IV PARTICIPANTS

The previous chapter described a methodology for analysing cognitive differences between competition law professionals who varied in their levels of specialist legal expertise. That methodology, which was aimed at highlighting readily identifiable and measurable differences of this kind, assumed the involvement of 15 to 20 appropriately qualified individuals, all of whom would participate as volunteers without any financial or other reward. These volunteers were to be ranked according to their levels of likely expertise, thereby enabling their placement within groups A, B and C, which were aligned to Hoffman's Scheme's categories of apprentice, journeyman, expert and master.

The previous chapter also described a testing framework constructed around a specific representative task and a knowledge-elicitation technique that would enable the recording of think-aloud verbal protocols for subsequent analysis. That framework required the involvement of participants whose various metrics and levels of likely expertise could be confirmed and compared, which in turn required a methodology that would provide a firm foundation for the subsequent tests and analyses.

This chapter documents the profiles of the participants who volunteered to take part in this study. This information includes background information for each participant as well as comparative and statistical data relevant to the discussion of cognitive differences in subsequent chapters. This chapter also documents the scoring and ranking process which culminated in the allocation of participants to Group A, Group B or Group C according to the format of Table 3.1 from the previous chapter.

Part A of this chapter recounts how approximately 80 lawyers and economists who specialise in competition law were invited to be involved in this study, the professional backgrounds of those 20 individuals who ultimately participated, and what their differences and similarities were with respect to a number of relevant criteria. This includes a comparative analysis of their years of general and specialist experience, their statuses within their firms or organisations, and the numbers of merger matters on which they had previously worked. Statistical data are also provided to establish dividing lines between those participants who were, for the purposes of the selection

process, presumed to have had greater expertise than other participants, which was a necessary requirement in the selection process and a reference for subsequent rankings.

Part B describes how the five previously described measures of likely expertise were applied to these 20 participants. This includes details concerning participants' sign-off authority (Measure 1), whether or not they satisfied the so-called 10,000-hour rule (Measure 2), evidence of their capabilities in relation to conceptual analysis (Measure 3), instances of exceptional reasoning (Measure 4), and comprehension errors when assessing the test cases (Measure 5). This information is summarised in tabular form and overall ranking scores determined according to the procedures described in the previous chapter.

The remainder of Part B considers statistically significant differences between participants in Group A, Group B and Group C, which groups were based on the previously explained assumptions regarding the range of participants in this study. The resulting comparisons are discussed in the context of the analysis to be undertaken in chapters 5 and 6.

Part C compares the expertise ranking of individual participants according to the methodology used in this study, with the rankings of leading professionals according to two international legal directories. This exercise provides further confirmation of the high-levels of expertise represented by higher-ranked participants. It also permits a comparative discussion of the differences and similarities between the results of this study's ranking methodology and these alternative ranking schemes.

A *Overview of Participants*

The process of selecting participants for this study began with sending standardised e-mails to partners and senior associates in major Australian law firms – and to their counterparts in economics firms and competition authorities – spread across all the main commercial centres in Australia. These e-mails explained the nature of the study and pro-actively sought to allay prospective participants' concerns about issues such as confidentiality and the required time-commitment to complete a test interview. Approximately 80 e-mails were sent out in a staged or staggered process so as to gauge and manage recipients' responses thereby ensuring, as far as practicable, an even balance between partner and non-partner level participants. After 20 volunteers had

agreed to participate and were tested on a rolling basis, no further e-mails were sent out.[378]

Of the 20 individuals (14 lawyers and six economists) who volunteered to participate in this study, 12 were from five of the top seven law firms in Australia as measured by total revenue for the 2011/2012 financial year.[379] Another participant was from a leading boutique firm which had for several years won national awards for its competition law practice.[380] The remaining legal professional, who had previously been a senior competition law partner in a national law firm, held a very senior position with a regulatory authority. Four of the economists who participated were from three of Australia's most prominent competition economics firms, that is, from firms that are part of a small group of economic consultancies routinely engaged by the largest Australian law firms and whose consultants consistently appear as expert economists in Australian competition law court cases.[381] The other two economists were from regulatory authorities. Geographically, at the time of their interviews, participants were spread across the major business centres in Australia, with eight located in Sydney, six in Melbourne, four in Brisbane, and one each in Perth and Canberra.[382]

By virtue of the selection methodology used, these 20 participants were divisible into two broad categories, namely, partner-level and non-partner-level. Within each category there were 7 lawyers and 3 economists. In the preliminary analyses of participants in the discussion below, the partner-level participants were numbered anonymously from S01 to S10 (lawyers were numbered S01 to S07 and economists S08 to S10). The non-partner participants (who all confirmed that they did not have

[378] Three participants, who were all partner-level specialists, were previously known to the researcher. Another four individuals volunteered to participate at the suggestion of other participants, who had either undertaken to contact the prospect directly or permitted the researcher to use their names by way of introduction. These additional volunteers comprised one partner-level participant and three non-partners.
[379] 'Top 35 Law Firms by Revenue for FY 2011/12,' in 'Australia's Top Law Firms Revealed,' Business Review Weekly, 1 August 2012 http://www.brw.com.au/p/sections/focus/australia_top_law_firms_revealed_vvZ5sZcs7mnCsgi2pdqCFM
[380] All the law firms represented in this study were also included in the list of 10 Australian law firms rated for their competition law expertise in *Who's Who Legal* and *Chambers and Partners* (www.whoswholegal.com and www.chambersandpartners.com accessed 10 May 2014).
[381] All three economics firms represented in this study were included in the list of six Australian economics firms specializing in competition economics published by *Who's Who Legal* (www.whoswholegal.com accessed 10 May 2014). *Chambers and Partners* does not rank competition economists.
[382] The study's sole specialist in New Zealand competition law was temporarily resident in Australia at the time of his interview.

sign-off responsibility within their organisations) were numbered S11 to S20 (lawyers numbered S11 to S17 and economists S18 to S20).

Statistically significant differences and similarities between these two groups of participants were identified based on their answers to a series of preliminary questions. These questions (reproduced in Appendix A) related primarily to participants' experiences and qualifications. Insofar as the two groups could be demonstrated to be significantly different in key areas (such as years of specialist experience), but similar in others (such as degree of specialisation and number of mergers considered in the last five years), the more likely there would be a broad allocation of participants across the expertise spectrum and therefore the more meaningful the ultimate results of the study. The same statistical methods and assumptions were applied to these tests as were applied to the tests for statistical significance reported elsewhere in this study.[383]

On this basis, it was possible to confirm with 95 per cent confidence that the partner-level participants, taken as a distinct sample group, had:

- more years of general professional experience than the study's non-partners (an average of 30 years compared with an average of 9.3 years);

- more specialist professional experience in competition law or competition economics (22.9 years compared with 7.9 years);

- worked on more merger cases during their careers (108.6 compared with 30.6); and

- more years of sign-off responsibility (16.9 years compared with 0 years).[384]

These were the key areas of difference identified during the preliminary assessment of participants.

[383] All tests for statistical difference in this study took the form of Student T-Tests with a null hypothesis of 'no difference between the two sample groups' and assumptions of a two-tailed distribution, two-sample unequal variance, and normal distribution. The confidence interval was set at $\alpha = 0.05$ (95% confidence level).

[384] Each non-partner was asked to explicitly confirm that they did not possess sign-off responsibility for their firm.

The following Table 4.1 records individualised information for each participant. To ensure anonymity, the numbering of participants has been randomised – in this instance only – so that, for instance, participant S06 is by his designation a partner-level lawyer (being within the group S01 to S07), but he is not the actual S06 whose verbal protocols are discussed in subsequent chapters. In other words, within each of the categories of lawyer partner, partner-level economist, non-partner lawyer and non-partner-level economist, the numbering has been randomised. But the placement of an individual within an expertise category is accurate. This is to avoid readers with knowledge of a particular participant's background (including the participant themselves) being able to trace their performance in this study from the data below.[385]

[385] Assurances of confidentiality and anonymity were critical to securing participants' agreement to be involved in this study. The fact that audio recordings would be made of participants' real-time analyses of cases they had never seen before, was something that several participants said they were uneasy about, but that given the privacy safeguards of the study, they could accept it. One participant said that the biggest issue for him was having a recording made of his initial reactions to problems, where those reactions would normally not be expressed. However, he also understood – as did the other participants – that this requirement went to the heart of the testing and analytical methodology of the study. It is also important to note that where a particular participant went poorly on a test case, it was preferable – as a matter of courtesy, if nothing else – that they not be able to identify themselves as that participant.

TABLE 4.1 – Participant Differences: Partner-level vs Non-partner Level

	Participant	General Experience (Years)	Specialist Experience (Years)	Sign-off Responsibility (Years)	Merger Matters Career to Date (Number) [386]
Partner-level Lawyers	S01	35	30	25	150+
	S02	51	25	25	150+
	S03	20	14	3	10
	S04	23	25	12	150+
	S05	18	13	10	120
	S06	23	10	14	6
	S07	35	35	25	150+
Partner-level Economists	S08	30	30	15	150+
	S09	35	35	20	150+
	S10	30	12	20	50
	Average (SD)	**30 (9.8)**	**22.9 (9.8)**	**16.9 (7.4)**	**108.6+ [387] (61.6)**
Non-partner Lawyers	S11	13	13	0	100
	S12	7	6	0	6
	S13	8	7	0	20
	S14	10	7	0	20
	S15	10	12	0	5
	S16	8	6	0	50
	S17	9	6	0	20
Non-partner Economists	S18	8	7	0	20
	S19	10	5	0	30
	S20	10	10	0	35
	Average (SD)	**9.3 (1.7)**	**7.9 (2.8)**	**0**	**30.6 (27.7)**

In terms of similarities between these two groups, it could not be shown that there was any statistically significant difference between the group of 10 partner-level participants and the group of 10 non-partner participants in relation to:

- their degree of specialisation in competition law during the last twelve months as measured by the percentage of their billable time spent on competition

[386] The number of mergers a participant had worked on during their career was a generalised point of comparison only given that many participants found it difficult to remember how many merger matters they had worked on and because their level of involvement and responsibility varied significantly. The nature of the matters themselves also made direct comparisons problematic. For example, one non-partner had worked on the same merger for two years, while another had considered 5 mergers, none of which had led to a merger review application.

[387] Several expert participants estimated that they had worked on over 200 merger matters during their careers. Rather than seek a specific number, the figure of 150+ was used in these instances. This was not considered a significant rounding issue given the already noted problems with comparing numbers of merger matters directly.

matters (the average across all participants was 77%);

- the number of merger matters they had worked on during either the preceding 12 months (3.8 matters on average) or over the last five years (20 matters on average);

- for those partners and non-partners who had worked as employees at a competition authority, the number of years of experience they had spent in those roles (the average for five partners and five non-partners with such experience was 8.2 years); and

- the number of years that these 10 participants had spent reviewing mergers at a competition authority (average 4.1 years).

In addition to these statistics, other comparisons were made on a simple arithmetic basis. These comparisons highlighted additional differences between the partners and the non-partners in this study. For instance, three of the former category of participants had taught in post-graduate competition law courses, but none of the latter had. This was consistent with the Hoffman Scheme's traditional definition of a master being one who is, among other things, qualified to teach lower-level specialists.

Further, as noted in the previous chapter regarding Hoffman's Scheme, being a master (the highest level of expert) meant 'being part of an elite group of experts whose judgments set the regulations, standards, or ideals'[388] within their domain. Four of the partner-level participants in this study ostensibly demonstrated these attributes having authored authoritative competition law reference-texts, appeared as expert witnesses in landmark court cases, set the policies of competition authorities, and influenced the formulation of competition laws themselves.

On the other hand, there were areas where no obvious arithmetic differences existed between the partners and non-partners involved in this study. The number of participants with post-graduate qualifications in competition law was equal between the two groups. Each group contained one participant with a PhD and two participants

[388] Robert R Hoffman, 'How Can Expertise be Defined? Implications of Research from Cognitive Psychology' in Robin Williams, Wendy Faulkner and James Fleck (Eds), *Exploring Expertise: Issues and Perspectives* (Macmillan, 1998) 84.

with a competition law LLM. Nine or 90% of the partner-level participants volunteered that merger review cases were a focus of their practices, while 8 or 80% of the non-partners made the same claim. Four or 40% of both groups stated that they spent more than half their time on litigation or 'back-end' work. Four partner-level participants and three non-partners stated that they devoted most of their billable time to advisory and transactional work, which is also termed 'front-end' work. The remaining two partners and three non-partners characterised their practices as a mixture of both front-end and back-end work. Five partners and five non-partners had either worked for or were currently working for a competition authority.

The fact that all participants worked for leading law or economics firms, or with competition authorities, ensured that their work experience was likely to have been of a comparable quality. As discussed in the previous chapter, smaller organisations were considered unlikely to maintain substantial specialist practices in competition law or be able to attract the highest level clients with the most complex competition law issues, unless they were boutique firms that specialised in competition law or competition economics.

Only senior associates and partners (or their non-law equivalents) at these organisations were invited to participate. This ensured that, as a practical matter, their minimum years of specialist experience in competition law would be no fewer than five years based on the common practice that promotion to senior associate typically requires several years of specialisation in a particular area of law or economics. In the other empirical studies of legal thinking skills reviewed in previous chapters, most if not all of these non-partners would have been considered experts in those expert and novice comparisons.[389]

[389] For instance, in the research reported by Colon-Navarro and Weinstein (and also by non-law scholars such as Lundeberg), measures of expertise were predominantly used in a loose, relative sense. While the experts in these studies were ostensibly more expert (or at least more experienced) than the novices and subexpert in the subject field of law, their expertise in a more absolute sense was open to question. Colon-Navarro's four experts had 3, 5, 13 and 14 years of specialist experience in immigration law (Fernando Colon-Navarro, 'Thinking Like a Lawyer: Expert-Novice Differences in Simulated Client Interviews' (1997) 21 *The Journal of the Legal Profession* 107, 121). Weinstein's three experts were simply described as 'outstanding practitioners' with no reference to how many years they had practiced social security disability law (Ian Weinstein, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1, 18). Lundeberg's ten experts, who were presumed to be experts in reading court cases, were either lawyers or law professors 'who had practiced law or at least taught two years' (Mary A Lundeberg, 'Metacognitive Aspects of

The decision not to involve participants with less than 5 years specialist experience in competition law was made for a number of reasons, several of which were noted in parts B and C of Chapter 3. First, this study was not intended to be a comparison between, on the one hand, novices who had had little or no exposure to competition law issues and, on the other hand, much more experienced competition specialists. Moreover, it was critical that technical legal knowledge should not be a distinguishing characteristic between participants, at least not with regards the chosen representative task. Second, the representative task itself required a relatively deep understanding of legal and procedural issues that could not be assumed of less experienced participants. The prospect of otherwise having to rely on a simpler or more straightforward task that novices could attempt risked leaving the more experienced legal specialists unchallenged, which would have been a serious flaw in the study's overall design.

As to the prospect of designating participants with up to 10-years' experience to be apprentices, this was considered to be within the traditional categorisation of apprentices having 'about one to 12 years' experience in the craft guilds,'[390] and their being individuals 'undergoing a programme of instruction beyond the introductory levels.'[391] In addition, the next step up to journeyman – that is, 'a person who can perform a day's labour unsupervised'[392] – was not a status that could be assumed for all participants. Some of the more experienced non-partner level participants did, however, state that they had effective autonomy up to the point of sign-off in most instances.

Attention was also given to ensuring that no single firm or organisation was represented by more than three study participants. This was made possible through the adoption of the staggered approach to soliciting prospective participants. Where this number had been reached for any one organisation, another would be approached. However, for the most part there seemed to be a natural limit of three individuals per firm choosing to volunteer.

Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407, 410).

[390] M T H Chi, 'Two Approaches to the Study of Experts' Characteristics,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 21, 22.

[391] Ibid.

[392] Ibid.

These statistics confirmed, at least in a general sense, that the participants in this study were sufficiently similar in some areas and sufficiently dissimilar in others to ensure that the first-order methodological requirements of the study were met. Prima facie, all participants were experienced competition law specialists working at the higher end of their profession either in private practice with a leading law or economics firm, or with a competition authority. Moreover, no single firm or city could have been considered over-represented.

Second, every participant had demonstrably adequate specialist knowledge of the relevant merger review laws and procedures. This was subsequently confirmed insofar as no participant needed or requested any technical references or guidance to assist in their analysis of the test cases. Their apparent familiarity with merger review laws and procedures was not a precondition to any participant's selection. It was nevertheless confirmation that merger cases were an appropriate focus for a study of specialist competition law expertise.

Third, these 20 volunteer participants were plausibly divisible into two distinct groups with statistically significant differences, namely, a group of 10 participants with partner-level status and a group of 10 participants without such status. From this arose a reasonable expectation – anticipated in the design of the study – that within the former group would likely be experts and masters within the meanings of these terms as used in Hoffman's Scheme, and within the latter group journeymen and apprentices. This mix of otherwise broadly homogenous specialist competition lawyers and economists was considered an appropriate starting point for further ranking.[393]

B *Participant Rankings*

The first five sections of this part record how the five measures of likely expertise discussed in Part C of the previous chapter were applied to each individual who volunteered to participate in this study. The measures used, as well as their associated scoring and weighting, were described at length in the previous chapter. The

[393] One distinction not considered relevant to the findings of the present study was the gender of the 20 participants who took part, of whom five were women and fifteen were men. There was no targeting of any one gender in the solicitation of volunteers. Participants' involvement was ultimately a matter of personal choice and availability. Regarding the use of a single pronoun, this was considered preferable to avoid unnecessarily highlighting the gender of participants throughout the thesis, and the male pronoun was confirmed via a literal toss of a coin.

discussion below is solely concerned with the application of the relevant principles and substantive requirements.

The last section of this part summarises the aggregated results of the scores of likely expertise for each participant. This includes comparisons between the three different groups of participants as allocated according to Hoffman's Scheme. This comparison includes additional statistical analysis to indicate the ways in which these groupings differed from and were similar to the initial grouping of participants as either partner-level or non-partner level legal specialists.

1 *Partnership (Sign-off Responsibility)*

For most participants in this study, their status as, on the one hand, law firm partners or principals in economics firms, or, on the other, non-partners or senior associates within their organisations, was clear. What was not clear for some members of the former group was whether or not their promotion to partnership or principal reflected their specialist competition law expertise at the time of their promotion.

Two participants had become competition law specialists less than five years prior to becoming partner (S03, S07). Another two had been partners for up to eight years before they became specialists in this area of law (S02, S08). Accordingly, these participants' promotion to a position with sign-off responsibility was ostensibly for reasons other than their competition law expertise.

Within the framework of the present assessment of specialist expertise, these four participants were assigned a neutral score on this particular measure, but subsequently received the maximum score under the 10,000-hour measure. Had they not held partnership status at the time of testing, such as the senior associates in this study, they would have received a -1 score under the present measure. As it was, they were scored at '0.'

In relation to the three participants who worked for competition authorities, it was initially assumed that those individuals who were likely to be made partners in a private sector firm, should they make the move, were to be scored positively on this measure. The first of these participants (S01) had previously been a senior partner in a leading national law firm. Moreover, he had originally been promoted to the partnership in that

firm as a competition law specialist. Accordingly, his status as an expert on this measure was clear, and he received a score of +1.

The second participant (S10) had held senior roles in competition authorities for well over a decade. His decisions were in most practical senses final in those capacities, though formally subject to official authorisation. During the participant-selection process he had been categorized as a partner-level participant owing to the level of responsibility he held within the government, which was ostensibly equivalent to that of a law firm partner. However, he was given a neutral score of '0' on the basis that there was no actual confirmation of him achieving partnership status as a result of a vote of his peers. In other words, there was no recorded attestation by those who stood to be disadvantaged by his judgments on competition law issues.

The third government employee (S19) had previously worked as a consultant with a prominent economics firm and had only recently re-joined the government. His relatively limited years of experience and confirmed status at below partner-level while in private practice, confirmed that he should be classified as a non-partner without sign-off responsibility and hence was given a score of -1.

These individual assessments against this first measure of specialist legal expertise are summarized in the following table.

TABLE 4.2 – Sign-off Responsibility Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Sign-off Responsibility	+1	0	0	+1	+1	+1	0	0	+1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

2 *10,000 Hours*

With the exception of one partner (S02), all participants with sign-off responsibility had twelve or more years' experience as competition law specialists and were scored at +1 on this measure. Participant S02 had ten years' experience, and according to the methodology described in the previous chapter, was scored at '0.'

Seven of the non-partners participants (S11, S12, S14, S15, S16, S19 and S20) had fewer than 8 years' experience as competition law specialists and were therefore given scores of -1. Two non-partners (S13 and S17) had twelve or more years' specialist

experience and were given scores of +1. Like participant S02, non-partner S18 had ten years' specialist experience and was scored at '0.'

The following table summarizes these results.

TABLE 4.3 – 10,000 Hours Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
10,000 hours (10 years)	+1	0	+1	+1	+1	+1	+1	+1	+1	+1	-1	-1	+1	-1	-1	-1	+1	0	-1	-1

3 *Conceptual Depth*

Partner-level participants S04, S05, S08 and S09 had each identified issues indicating that their analysis had been conceptually deep either in their assessment of Case A or Case B, or of both cases. Non-partners S15 and S19 performed similarly well on these cases. These six participants were therefore each scored at +2 according to the scale of conceptual depth established in the previous chapter. Which is to say, these participants had focused on product functionality and production substitution in relation to the concrete products of Case A and/or had highlighted the risk of vertical foreclosure in shipping services as a possible post-merger scenario in Case B.

Partner S02 was assessed as being at the next level of analysis and was scored at +1 on this measure. This was based on his having concluded in Case A that manufacturers of other concrete products could switch to produce the products supplied by the merger parties. Participant S17 received the same score for his analysis of demand-side substitution in Case A.

Partner S03 and non-partners S11, S16 and S18 undertook a balanced analysis of Case A and, with the exception of S18 who did not consider this second case because of his prior knowledge of the transaction, adopted a similar approach in Case B. Each of these participants was scored at '0' for conceptual depth. Non-partners S01 and S06 considered neither case and were therefore also scored at '0' by default.

Partner S10 was scored at -1 for conceptual depth because he focused mostly on geographic issues in Case A and did not identify the vertical issue in Case B. However, he did not go as far as non-partners S13 and S14, who both favoured arguments in these cases as to likely limited competition between the parties pre-merger. These arguments

were aimed at allaying post-merger competition concerns, but were based on superficial (and erroneous) assumptions about the relevant industries. Accordingly, both these participants were scored at -2 on this measure.

Non-partner S12 focused briefly on demand-side substitution in Case A and, while erroneously concluding that the merger parties in Case B serviced different customers, did not go as far as S13 and S14 in relying on 'no-overlap' arguments. Accordingly, like S10, this participant was scored at -1.

Non-partner S20 was amongst the most superficial of all participants in his analysis of both cases, concluding in Case A that because he had not personally heard of either party they were likely to be small players in the concrete market and that therefore there was unlikely to be any serious competition issue. In Case B, he concluded that that transaction would be problematic principally because of the substantial size of the merger parties' businesses in the national marketplace for freight-forwarding services. This participant was scored at -2 for conceptual depth.

These scores are recorded in the following table.

TABLE 4.4 – Conceptual Depth Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Conceptual Depth	0	+2	0	+2	+2	0	0	+2	+2	-1	0	-1	-2	-2	+2	0	+1	0	+2	-2

4 *Exceptional Reasoning*

There were six examples of exceptional reasoning across all study participants, that is, reasoning that was unique in the context of a given case and which was material to determining how the case was assessed by the participant in question. For each of these examples, an assessment was made against the ideal response described in Section 4 of Part C of the previous chapter.

Partner S01, because of his status as a specialist in New Zealand competition law, considered cases that no other participant considered. It was therefore difficult to determine the uniqueness of this participant's assessments. However, his approach was considered sufficiently similar in tone to the reasoning of S06 in Case C (as described

below), that his reasoning was assessed as being at least as exceptional and deserving of a positive score against the ideal response.

The second case considered by participant S01 (Case F) concerned a merger transaction involving book publishing businesses. Within less than 10 seconds of commencing to read the test case document, he had concluded that the case would be cleared by the CC. The basis for this conclusion, as he went on to explain, was that publishing was a 'declining business'[394] with 'low barriers to entry.'[395] After re-confirming that clearance was 'not a problem'[396] and that it was therefore not necessary for him to read the case any further,[397] participant S01 stated that the only remaining task would be to 'plod through and explain it all'[398] to CC staff.

This assessment was scored at +2 on the grounds that participant S01 had relied on his reasoning strategy to a material extent, that his reasoning was based on known facts, experience and inferential insight, that it reflected extensive specialist knowledge, and that it was novel, creative and compelling. It also turned out to be correct inasmuch as the transaction was eventually cleared by the CC. However, being correct about the final outcome of a case was not necessarily confirmation of specialist expertise, as was apparent in the exceptionally superficial – but ultimately correct – reasoning demonstrated by non-partner S20 in Case A as discussed further below.[399]

Participant S03 provided a very confident assessment in his analysis of Case C. His confidence was based on an assessment that the ACCC's questions were 'of a very general nature'[400] and looked like 'they're designed to assuage some complaint'[401] rather than because a substantive issue existed. Moreover, he noted that the ACCC's inquiries were being made 'some time after the fact,'[402] which further suggested a lack

[394] S01F Line 7.
[395] S01F Line 25.
[396] S01F Line 27.
[397] S01F Line 29.
[398] S01F Line 36.
[399] Given the lack of information provided in the test-case documentation, it was considered unrealistic to expect participants to offer conclusive opinions. At least, it was not something that was considered would be so common as to be a reliable reference for ranking purposes where ideally all participants would have given conclusive opinions which could have been assessed for accuracy. That said, it remained feasible to include the accuracy of the firm opinions offered by some participants as one of the comparative differences identified in subsequent phases of the study. In which case, not including such a measure here in the ranking phase would avoid a potential circularity issue.
[400] S03C Line 99.
[401] S03C Line 101.
[402] S03C Line 103.

of urgency or concern. It is to be recalled that Case C concerned an already completed transaction.

But the ultimate confirmation that the transaction would be cleared (or not challenged by the ACCC by way of a court action seeking divestiture) was S03's observation that as an international merger it was likely that formal opposition could only be initiated under Section 50A of the *Competition and Consumer Act* 2010. This section, he noted, 'creates another whole level of protection for the entities because the Commission can't actually do that much about it if it falls under this statutory provision,'[403] which – in the broadest terms – requires more complicated administrative steps than initiating court action under Section 50 as stated in the ACCC's letter.

S03 concluded his assessment stating, 'so I would be advising the parties in half an hour that they really don't have anything to worry about, subject to finding out some more information.'[404]

This assessment was scored at +1. First, it was unique (no other participant had cited the relevance of Section 50A), it was based on known facts and on insight consistent with a high degree of specialist legal knowledge, and it was sufficiently compelling to convince S03, not unreasonably, that it would provide an adequate answer to the ultimate question regarding the likelihood of regulatory opposition. However, as a matter of statutory interpretation, it is doubtful that Section 50A would actually apply in this instance given that Section 50, as asserted by the ACCC, would likely oust the former provision's jurisdiction. Moreover, Section 50A has not to date been used by the ACCC in any case, notwithstanding it has been part of Australian competition law since 1986.[405]

Accordingly, S03's risk assessment strategy satisfied to a large extent the elements of an ideal response as specified in this study. It did not, however, rank as highly as the responses of the other positively-rated exceptional responses by S01, S06 and S09

[403] S03C Line 113-114.
[404] S03C Line 116-117.
[405] Fels, Henrick and Taylor convincingly argue that this 'curious provision' is only likely to apply in 'very limited circumstances.' This is because by virtue of Section 50A(7), Section 50A may only apply where Section 50 does not (but not the reverse). This creates a formidable subject-matter jurisdiction hurdle to Section 50A's application in all but the rarest cases: Allan Fels, Sharon Henrick and Martyn Taylor, 'Australia,' in Maher Dabbah and Paul Lasok (eds) *Merger Control Worldwide* (Cambridge University Press, 2012) 78, 83.

(which were in technical legal terms completely accurate) and therefore was given a +1 score rather than a +2 score as these other participants were given.

Participant S06 sought merely to confirm, as a matter of fact, that the transaction in Case C had been concluded or 'signed, sealed and delivered' to use this participant's own words.[406] Having satisfied himself that this was the case, he concluded within less than 30 seconds that the only possible remedy was divestiture and that in his view 'the chances of a court granting divestiture of any merger case in Australia – unless it's such a clear cut case – is going to be virtually zero.'[407] He then went on to recall two past cases which were relevant to this point.

Like S01, participant S06 very quickly formed a concluded and compelling view on the likely outcome of the case he was considering. Whereas several other participants spent over ten minutes before making their assessment in Case C (three participants – S12, S18 and S19 – took 14 minutes and longer), this partner-level participant took just under two minutes to establish perhaps the strongest conceptual argument based on the limited information available. Rather than provide a direct answer based on the substance of the case, S06 framed the central issue as whether or not this was likely to be an extreme case – more extreme than any other in Australian competition law history – such that it might result in the first ever successful divestiture action by the ACCC.

Participant S12 had also vocalized that the ACCC had never before brought a successful divestiture case against a consummated merger. However, his immediate follow up thought was 'but who's to say this is not going to be the first.'[408] Instead of realizing immediately at that point, as S06 had done, that Case C would therefore need to be the most extreme such case to be in jeopardy, S12 then embarked on a 14 minutes and 30 seconds analysis in which he first concluded that the transaction was likely to be opposed. Subsequently he said the case would likely be cleared, but that work would need to be done to overcome what he had erroneously considered was 'a bit of confusion here on the ACCC's part'[409] about the precise nature of the transaction itself.

[406] S06C Line 11.
[407] S06C Line 17-18.
[408] S12C Line 33.
[409] S12C Line 315.

Participant S06's reasoning in Case C was considered both unique and compelling. It also drew on his deep specialist knowledge of this area of law (including a recital of the details of two previous cases, both of which were decided almost a quarter of a century earlier[410]), and reflected a creative approach in that it established a narrow focus for further analysis, that is, whether or not this case was likely to be the most extreme in anti-competitive terms compared to all previous cases the ACCC had considered since the commencement of the relevant merger laws in the late 1970s. Placed in that context, the case became immediately tractable, subject to confirming a limited number of factual details. Participant S06 was therefore scored at +2 for this example of exceptional reasoning.

Participant S09 demonstrated exceptional reasoning in both the cases he considered. In respect to Case A, he was the only participant to identify that there could be production substitution possibilities such that a manufacturer of reinforced concrete box culverts could switch to producing concrete drainage pits, depending on relative price and demand signals. This insight suggested a basis for viewing these products as co-existing within the boundaries of the same product market. The analytical effect of this perspective was to broaden the parameters of the relevant market and reduce competition law risk.

In Case D, participant S09 suggested that the best way of determining the geographic extent to which individual rural merchandise stores supplied their products (a critical issue in the case) was to review invoices held in each store's internal records and determine from them the locations of their customers (and what they bought). This, he opined, was a far better approach to determining how far farmers were willing to travel to purchase their rural merchandise than to ask the farmers directly, which he believed demonstrated a lack of understanding on the part of ACCC staff.

Both of the above perspectives were unique. They also indicated S09's deep expert knowledge and practical experience in competition law matters. Moreover, they raised issues on which these cases would be likely to turn in evidentiary and analytical senses, although S09 admitted to not knowing the outcome of his suggested further inquiries. As both assessments were insightful, compelling and relevant to determining how easy

[410] *Trade Practices Commission v Australian Meat Holdings Pty Ltd* (1988) 83 ALR 299 and *Re Arnotts Limited v Trade Practices Commission* [1990] FCA 473.

clearance was likely to be in these cases, this participant's exceptional reasoning was scored at +2.

Participant S12 has already been noted for not linking the significance of Case C being concerned with a concluded merger and the fact that the ACCC had never successfully secured divestiture in such cases. Compared with S06's more insightful assessment of those facts, S12 arguably performed less like an expert in that instance. However, S12 had done better than most other participants who had not even recalled the historical track record of the ACCC in such cases.

Yet it was in his assessment of Case B where S12 most clearly demonstrated his less-than-expert status. He was the only participant to misunderstand the relationship between the acquiring party in that case and its existing joint-venture shipping partner. At one point he correctly identified the vertical integration between Toll (the freight-forwarder) and Toll ANL (the shipping joint-venture). However, he then became convinced that Toll ANL was also a freight-forwarder in its own right and a direct competitor of Toll and the target business, Linfox Trans-Bass. This was incorrect factually and misleading in the sense that it led S12 to characterize the merger as a reduction in the number of competitors in the market from three to two. This, in turn, led him to conclude that his 'competition advice would be that this [merger between Toll and Linfox Trans-Bass] is likely to be opposed by the ACCC.'[411]

As no other participant relied on a similar series of mistaken inferences, and because this line of reasoning was based on factual errors that ultimately led S12 to conclude that the case would be opposed – which it was not – this example of exceptional reasoning was scored at -2 against the ideal response.

The last example of exceptional reasoning was that of non-partner S20 in Case A. This participant's conclusion that there would be no competition issue in this case was based on the following analysis:

> If I had to say [how easy clearance would be], I guess I think on the one hand … concrete products sort of have a history of competition issues … on the other hand I

[411] S12B Line 238-239.

> haven't heard of these companies … so I'm assuming they are … perhaps fairly small ones … so there is a fairly good chance that the merger would go, will be cleared.[412]

While the size of the merger parties was noted as a relevant factor by other participants in other cases (S20 again noted this factor in Case D), only S20 raised this issue in this case. Moreover, he relied substantially if not wholly on this factor, doing so even while acknowledging that the only basis for his inference about the size of the parties was that he had personally not heard of them. Rather than try to understand the significance of the parties' businesses based on the information provided about how many manufacturing and sales sites they had established (the acquirer had a presence in most Australian states), who their customers were, and the degree of product substitution that might exist (both on the demand and supply-sides of the market), this participant relied on his personal experience in an industry with which he admitted he was unfamiliar.

This reasoning strategy was scored at -2 on the basis that it relied on dubious assumptions, did not reflect an extensive, specialist knowledge in competition law, and, while novel, was not compelling.

The other participants in this study were scored at '0' on this measure, not because their reasoning strategies were necessarily better than those of S12 or S20 or worse than those of S01, S03, S06 or S09, but because their approaches were more common. The sole purpose of this measure was to identify outlier reasoning strategies which were both unique and material to the assessments of the participants in question.

The results against this fourth measure of likely expertise are summarized in the following table.

TABLE 4.5 – Exceptional Reasoning Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Exceptional Reasoning	+2	0	+1	0	0	+2	0	0	+2	0	0	-2	0	0	0	0	0	0	0	-2

[412] S20A Line 93-103.

5 *Comprehension Errors*

The most common error made by participants in this study concerned the ACCC's description of the merger parties in Case B, although comprehension errors were identified in several other cases as well. In Case B, the ACCC's letter had provided a description of the cargo carried by the shipping joint-venture in which the acquirer, Toll, was a party. This was followed by a description of the freight-forwarding service provided by the target, Linfox Trans-Bass. The result was that several participants mistakenly assumed that these businesses had different customer bases: The former as a carrier of trucks, motor vehicles and infrastructure equipment, and the latter as a carrier of consumer goods and retail products. In fact, both Toll and Linfox Trans-Bass provided the same kinds of freight-forwarding services, but Toll was the only one that was vertically integrated into shipping.

Partners S02, S04 and S10, and non-partners S11, S12, S13 and S14, considered that *both* parties were integrated logistics operators. This was incorrect as only Toll was vertically integrated over the shipping routes in question. Partner S07 and non-partners S16 and S20 did not consider the vertical issue at all, but rather proceeded on the basis that the transaction only concerned freight forwarding and, therefore, involved only a horizontal market-concentration issue. Each of these participants was scored, initially at least, at -1 according to the scoring scale described in the previous chapter.

The reason for this score of -1 rather than -2 was that while errors of comprehension were made by these participants, their assessments were not dependent on their mistaken understanding of the vertical integration issue.

Participants S12, S13 and S14, on the other hand, made a further material error by concluding that the merger parties serviced different customers. But only S13 and S14 relied on this assessment to develop an argument that the parties' operations were therefore more complementary than competitive, and hence there was unlikely to be a competition problem. This was a material error by these two participants and their scores were further reduced to -2 on this basis.[413]

[413] A '-2' score for S13 was further confirmed by his analysis of Case A, where he mistakenly considered drainage pits to be functional substitutes for reinforced concrete box culverts. This was a material error because he subsequently concluded, incorrectly, that barriers to entry would be lower insofar as simply being able to dig a pit was substitutable with making a concrete box culvert.

Participants S10, S12 and S17 noted a possible difference in the customer bases of the merger parties, but did not make anything more of this issue. Participants S10 and S17 had identified other more compelling issues. Participant S12, on the other hand, dismissed the no-overlap argument on the basis that over time the parties could become competitors, even if they were not competitors now. Accordingly, he would have presumably relied on this argument if time had not been an issue. In any event, on this specific issue this participant made an ostensibly less material error than S13 and S14, and for this reason was scored at -1 in relation to his assessment of this case.

Participant S12 was, however, the overall worst performer of all participants in terms of material comprehension errors. As previously noted, he was the only participant to erroneously view the shipping joint-venture Toll ANL as an independent competitor in the freight-forwarding market. He also mistakenly recalled that Toll had bought ANL 'a few years ago,' according to an article he believed he had read previously.[414] In Case C he contended that the ACCC had made a typographical error in its market inquiries letter insofar as it had confused the 'fact' that only one of the parties hired intermediate bulk containers while the other supplied them for sale. The reality was that both parties hired them and the ACCC's letter was correct. In Case D, S12 mistakenly believed that the merger parties sold tractors to farmers, that the ANZ Bank had previously bought the target companies' banking operations, that the target company provided a form of consumer credit with interest free periods like a retail furniture store, and that the acquirer 'made' its own fertilizer.

While not all of these comprehension errors could be considered material, some were, and the others implied poor recall and self-monitoring skills. Accordingly, S12 was scored at -2 for multiple errors, some of which were material.

The best performing participants under this measure were partners S05 and S08, and non-partners S15 and S19. Participants S05 and S08 made no apparent errors in any test cases and despite the ACCC's poorly drafted description of the parties had managed to identify not only that Toll was the only vertically integrated party in Case B, but that vertical foreclosure was likely to be an issue in the case. Participants S15 and S19 had also correctly identified both the vertical integration of Toll alone and the

[414] S12B Line 96. There is no record of Toll having purchased ANL. Rather, the ANL website (www.anl.com.au) describes the company as being 'part of the CMA CGM Group, the third largest shipping line in the world.'

vertical foreclosure issue. However, these two non-partners also made one error each. In Case D, S15 had thought that 'crop protection' was a financial service provided by the parties when in fact it involved the supply of chemicals and pesticides. In Case C, S19 mistakenly observed that this case was the first time that a countervailing power issue had been raised in the test cases to that point. In fact it had been previously mentioned in Case A. On balance, these were considered trivial errors that should not affect the scores given to these participants.

Accordingly, participants S05, S08, S15 and S19 were the only participants in this study to be scored at +1 on this measure. The remaining participants (S01, S03, S06, S09 and S18 who did not consider Case B and who otherwise made no apparent errors) were scored at '0,' which may be considered a conservative result given that at least some of them may have performed as well as S05, S08, S15 and S19 in terms of overcoming the ACCC's unintended 'red herring' about the differences in services provided by the freight-forwarding merger proponents in Case B.

The above results are presented in the following table.

TABLE 4.6 – Comprehension Errors Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Comprehension Errors	0	-1	0	-1	+1	0	-1	+1	0	-1	-1	-2	-2	-2	+1	-1	-1	0	+1	-1

6 *Summary*

The measure-by-measure scores described above as well as the cumulative scores of likely expertise for each participant are summarized in Table 4.7 below.

TABLE 4.7 – Summary of Likely Expertise Scores

MEASURE	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Sign-off Responsibility	+1	0	0	+1	+1	+1	0	0	+1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
10,000 hours (10 years)	+1	0	+1	+1	+1	+1	+1	+1	+1	+1	-1	-1	+1	-1	-1	-1	+1	0	-1	-1
Conceptual Depth	0	+2	0	+2	+2	0	0	+2	+2	-1	0	-1	-2	-2	+2	0	+1	0	+2	-2
Exceptional Reasoning	+2	0	+1	0	0	+2	0	0	+2	0	0	-2	0	0	0	0	0	0	0	-2
Comprehension Errors	0	-1	0	-1	+1	0	-1	+1	0	-1	-1	-2	-2	-2	+1	-1	-1	0	+1	-1
TOTALS	*+4*	*+1*	*+2*	*+3*	*+5*	*+4*	*0*	*+4*	*+6*	*-1*	*-3*	*-7*	*-4*	*-6*	*+1*	*-3*	*0*	*-1*	*+1*	*-7*

For the purpose of the analysis described in the next part of this chapter, a more useful presentation of this information (using just the total likely expertise scores for each participant), is reconfigured in Table 4.8 further below. This table is an adaption of Table 3.1 from Chapter 3. It clarifies which participants were within which of groups A, B and C. The non-partners identified during the participant-selection phase of this study are shown in the shaded boxes. When contrasted with the partner-level participants in unshaded boxes, the points of overlap can be seen amongst the Group B participants, who were categorized as an indeterminate mixture of all levels of likely expertise under Hoffman's Scheme.

TABLE 4.8 – Participant Groupings According to Scores of Likely Expertise

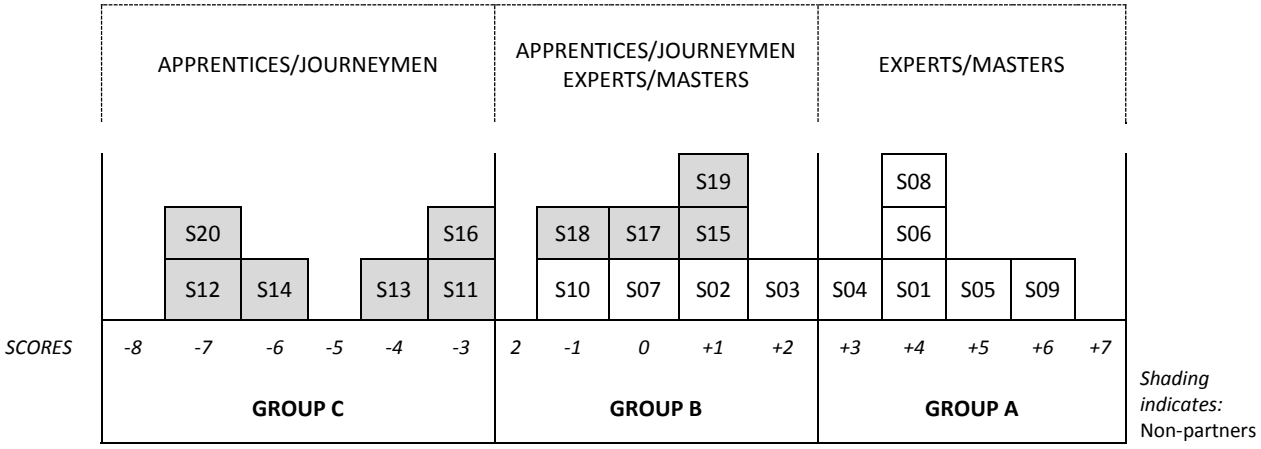


Table 4.8 shows that within Group B, non-partner participants S15, S17, S18 and S19 were in expertise terms indistinguishable from partner-level participants S02, S03, S07

and S10. In Group A, partners S01, S04, S05, S06, S08 and S09 appear towards the top (right-hand side) of the table as the most likely experts and masters in the study, while in Group C non-partners S11, S12, S13, S14, S16 and S20 appear towards the bottom (left-hand side) of the table as most likely apprentices and journeymen according to Hoffman's Scheme.

Within each of these three groups, the variation in scores was an important analytical feature. With a maximum range of 5 points between members of the same group (for example, Group B extends from -2 to +2, inclusive) there was reduced reliance on any one measure of likely expertise and thus an allowance for possible errors in the scoring process. By treating Group A participants as equivalent in terms of their levels of likely expertise, whether any individual in this group scored better or worse on any single measure (for instance, participant S09 received a +2 score on exceptional reasoning, while S04 received a score of '0' simply because he provided no evidence of exceptional reasoning) was less important. At the same time, members of Group A as a whole were clearly possessed of greater levels of likely expertise than members of Group C.

Testing for statistically significant differences between groups A, B and C yielded similar results to those generated from the tests for statistical differences undertaken between partner-level participants and non-partners as described in Part A above. While the sample sizes were smaller after dividing participants into three groups and therefore comparisons were potentially less meaningful, the following conclusions are noted as regards the points of difference and similarity between the participants in each group.

There were statistically significant differences between groups A and C in terms of:

- average years of general experience (Group A, 32.3 years; Group C, 8.7 years);

- average years of experience as a competition law specialist (Group A, 24.3 years; Group C, 7.3 years);

- average number of merger matters over their careers to date (Group A, 110 matters; Group C, 20.2 matters); and

- average number of years of sign-off responsibility (Group A, 17.5 years; Group C, 0 years).

These averages were approximately equivalent to the averages noted previously for partners and non-partners.

At the same time, there were no statistically significant differences between these two groups in terms of:

- average degree of competition law specialisation during the prior 12 months;

- average number of mergers matters considered during the last 12 months or last 5 years;

- average years working for a competition law authority; or

- average years working in the mergers branch of a competition law authority.

As to straight arithmetic comparisons, both groups had two participants whose practices were 'front-end' focused and two who maintained a 50/50 balance between front-end and back-end work. In addition, two members of Group A had a 'back-end' or litigation focus, as did two members of Group C. Both groups had one participant with a Master of Laws degree in competition law and one participant with a doctorate in competition law economics. Both groups had two participants with economics qualifications. Group A had four participants with legal qualifications. Group C had five participants with legal qualifications, one of whom had both economics and law qualifications but practiced as a lawyer rather than as an economist.

Comparing Group B with each of Group A and Group C, participants in Group A had on average more general professional experience than participants in Group B (32.3 years; 18.4 years). There was also a difference between average years of specialist experience in competition law, but only at the 90 per cent confidence level, which in this study was below the required threshold.

Compared to Group C, participants in Group B had on average more general professional experience than participants in Group C (18.4 years; 8.7 years). They had also worked on more merger matters (on average) during their careers to date (76 matters compared to 20.2 matters), and they had more years of sign-off responsibility, with no participant in Group C having had any sign-off responsibility.

One explanation for the above differences is that experience (and therefore, by extension, age) was the underlying factor separating these participant groupings. Simply by having spent more years practicing professionally, participants were more likely to have considered more merger matters and to have had more years as partners.

Apart from these areas, however, there were no other statistically significant differences between these groups, at least not across the range of statistics recorded in the preliminary questionnaire completed by all participants. Moreover, when estimating the age of those participants in Group A (age was not explicitly covered in the questionnaire), the average was likely to have been around 60 years. The only aberration in this sense was one participant in Group A who was estimated to be in his mid-40s.

The sole statistically significant difference between Group B and Group A participants was that the latter participants had, on average, more years of general professional experience, which simply reflects the fact that they were on average older. In terms of all other recorded data, there were no statistically significant differences between these groups, apart from Group A's participants being assessed as having greater levels of likely expertise.

As between groups B and C, the transition from the lower group to the higher group was associated with differences in years of general professional experience, number of mergers worked on career to date, and years of sign-off responsibility. Again, age can plausibly explain these differences.

A tentative conclusion about participants' levels of specialist legal expertise was therefore that older specialist lawyers and economists are likely to be more expert than younger specialists, and that this association continues to be observable as an individual's age rises to 60 years and beyond. In this sense, specialist legal expertise is

arguably cumulative and continuously increasing, even up to one's mid-70s, which was a milestone that had been reached by at least one of the participants in Group A.

Comparisons were also performed between Group A and the original group of partner-level participants, and between Group C and the original group of non-partner participants. No statistically significant differences were observed in either instance with regard to any of the above measures.

C *Comparisons with Industry Ratings*

Two of the more prominent research and ranking organizations that publish directories on specialist legal practices globally are *Who's Who Legal*[415]and *Chambers and Partners*.[416] Their directories, both in hard-copy and on-line formats, rank law firms and lawyers (and in the case of *Who's Who Legal* economics firms and economists as well) who provide Australian competition law advice. In broad terms, the research methodologies used by both directories are similar with some reliance on background information submitted by the subject law and economics firms, but with an emphasis on interviews conducted directly with peers and clients.

Who's Who Legal emphasizes that paid entries are not available in its directory and that its identification of legal specialists is 'based upon comprehensive, independent survey work with both general counsel and private practice lawyers.'[417] It further claims that its listings reflect 'specialists who have met independent international research criteria.'[418] *Chambers and Partners* similarly states that its ranking judgments are based on 'interviews with those active in the market – mainly clients (who can be law firms instructing other law firms) and other lawyers with whom they work' and on 'assessing recent work done.'[419] For firms and individual lawyers, the relevant criteria are 'legal knowledge and experience, their ability, their effectiveness, and their client

[415] www.whoswholegal.com (accessed 14 May 2014) describes itself as identifying 'the foremost legal practitioners in 34 areas of business law. We feature over 16,000 of the world's leading private practice lawyers in over 100 countries and pride ourselves on the integrity and authority of our findings. It is impossible to buy entry into this publication.'
[416] www.chambersandpartners.com (accessed 14 May 2014) describes *The Chambers Guides* as 'ranking the best law firms and lawyers since 1990' in over 185 jurisdictions. According to the website, a survey of 20,000 in-house counsel revealed that over half use the Guides to identify prospective legal service providers. This statistic is claimed to be higher than for any rival guides or law firm directories.
[417] www.whoswholegal.com/about accessed 14 May 2014.
[418] www.whoswholegal.com/about accessed 14 May 2014.
[419] http://www.chambersandpartners.com/methodology accessed 14 May 2014.

service.'[420] In its FAQ page, *Chambers and Partners* reiterates that a law firm cannot be guaranteed a listing in its directories merely by submitting information about its practices, and that listing ultimately relies on 'the feedback of clients, peers and other independent market sources.'[421]

Another organization, *The Legal 500*,[422] also provides directory services for users of specialist legal services, but appears to emphasise and rely to a greater extent on law firms furnishing the information upon which *The Legal 500* research staff base their assessments.[423] Primarily for this reason it was considered to be less objective and, in the context of the present study, less relevant than the ranking approaches of the other two directory publishers. In addition, whereas both *Who's Who Legal* and *Chambers and Partners* included some or all of the participants in this study in their listings of leading Australian competition lawyers and economists, none of these participants were included in *The Legal 500's* list of leading individuals in its competition and trade category.[424] This is notwithstanding that all the law firms represented in this study were listed among the top eight Australian firms ranked in that category by *The Legal 500*.[425]

Comparing the rankings of participants undertaken in this study with the rankings conducted annually by *Who's Who Legal* and *Chambers and Partners* served a two-

[420] http://www.chambersandpartners.com/methodology accessed 14 May 2014.
[421] http://www.chambersandpartners.com/faqs accessed 14 May 2014.
[422] *The Legal 500* describes itself as providing 'the most comprehensive worldwide coverage currently available on legal service providers, in over 100 countries.' It also claims that its law firm directory service is 'widely chosen [by its clients] for its definitive judgement of law firm capabilities over publications such as those compiled by Chambers and Partners or Martindale-Hubbell.' www.legal500.com accessed 14 May 2014.
[423] Whereas *Who's Who Legal* emphasizes that listings in its directory cannot be paid for and *Chambers and Partners* emphasizes the importance of its 150 full-time research staff whose primary role is to interview clients and other third parties, *The Legal 500's* editorial guidelines begin with a focus on timetables for law firms to submit information about their practice areas and to nominate referees. Firm nominated referees are explicitly given less weight by *Chambers and Partners* which explains its preference for (and the necessity of) independent third-party inputs, but these referees appear to be the primary source of third-party information for *The Legal 500*. In its website's FAQ page, *The Legal 500* states that its 'editorial research is based on firm's written editorial submissions, and follow-up research with firms and their referees.' Firms can also make submissions in relation to 'each practice area for which you believe your firm warrants recommendation.' Paid profiles for individual lawyers are also available within *The Legal 500*, and these must be written by the firms themselves. However, *The Legal 500* states that this information is not provided to the research side of its operations, which is described as separate from its paid services division.
[424] http://www.legal500.com/c/australia/competition-and-trade accessed 14 May 2014.
[425] Ibid.

fold purpose.[426] First, it confirmed the quality of the highest ranked participants in this study. While all the firms represented in this study were rated by these two industry research organisations as having the leading competition law departments in Australia, all partner-level lawyers in this study were also individually rated as experts by one or both of these publishers.[427]

Secondly, two of the 24 corporate lawyers identified by *Who's Who Legal* as experts in Australian competition law were participants in this study – and two of the 11 economics experts identified in that publication also participated in this study. While these absolute numbers are small, the selection process devised for this study yielded a not insignificant sample size of 8% of *Who's Who Legal's* expert competition lawyers and a sample size of 18% for the publication's expert competition law economists (the overall sample size, inclusive of both lawyers and economists, was 11% of the total population of national competition law and economics experts ranked by this publication). In relation to *Chambers and Partners'* rankings, only lawyers were eligible for inclusion in its list of 46 individual expert Australian competition law experts. Six of these lawyers participated in this study, which constituted a 13% sample from that list of the purported best competition lawyers in Australia.

Comparing the rankings in this study with the experts identified in the above guides to Australian competition lawyers and economists enabled the researcher to consider the extent of agreement between these peer and client-survey rankings, on the one hand, and rankings based on the methodology used here, on the other.

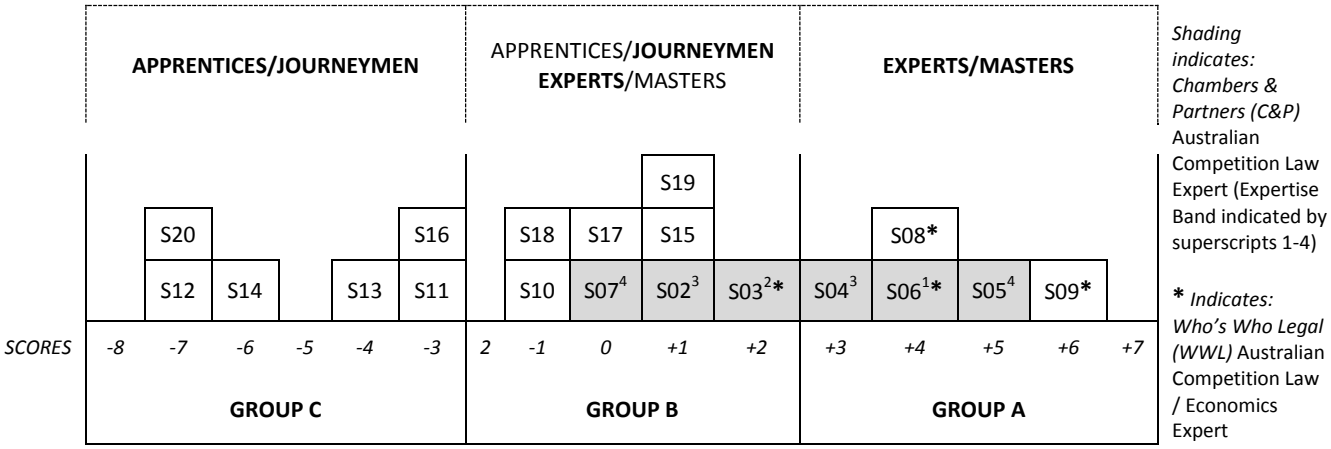
As can be seen from the following table, the participants in this study who were listed in *Who's Who Legal* were scored at no lower than +2 points (just outside Group A), with the majority being within Group A, which was the group assumed to include experts and masters in competition law and economics. Given that S09 and S08 were excluded from *Chambers and Partners'* assessment because they were not lawyers, it can be seen in the table below that that directory's assessment also coincided substantially with the ranking of lawyers according to their scores of likely expertise.

[426] No ranking publication was consulted during the scoring and ranking phase of participants in this study. It was only after ranking had been completed that confirmatory reference was sought from external directories.

[427] The only partner-level competition lawyer in this study not listed in either of these industry ratings publications is S01, who was a New Zealand competition lawyer (and hence would not appear in the Australian listings) and who is in any event no longer in private practice having taken up a senior position with a competition authority.

For the most part, the bands of experts used by *Chambers and Partners* comport with this study's rankings, albeit that participant S05 is ranked materially higher in this study (Inclusion in Band 1 of the *Chambers and Partners* framework denotes higher expertise than inclusion in Band 4).

TABLE 4.9 – Comparisons with Industry Ratings[428]



Significantly, the present study only sought to assess technical expertise, not whether an individual provided high or low levels of client service or was easy or difficult to work with in a team. Nor did it consider subjective inputs from third parties such as clients, colleagues or other industry players. In these respects the present approach contrasted with the methodological approaches of *Who's Who Legal* and *Chambers and Partners*, which both purported to give most weight to interviews conducted by their respective research staff with lawyers and law firm clients.

While this study referenced participants' total years of professional experience (both general and specialist) and partnership status, it relied more on three performance-based measures as assessed via participants think-aloud transcripts recorded during their analysis of test cases. This may explain why participant S05 (who had been a partner less than 5 years at the time of his interview) was rated more highly in this study than he was by *Chambers and Partners*. He simply had less time than other more established partner-level participants to develop a reputation amongst client and peers.

[428] Participant S01 is not included in this table because he is not an expert in Australian competition law, but rather in New Zealand Competition law, and hence would not be listed in either of these publications' Australian law assessments.

As it was, this participant was expressly singled-out in *Chambers Asia-Pacific Commentary* for his reputed 'excellent technical knowledge.'[429]

With reference to the differences between groups A, B and C, all four of the lawyer participants in Group A appeared in the *Chambers and Partners* 2014 list of 46 Australian competition law experts. At the same time, one of the five participants in this group was listed amongst *Who's Who Legal's* 24 legal experts, and two were included in its list of 11 competition economics experts. Three of the participants in Group B were also listed in the *Chambers and Partners* 2014 rankings, whereas one Group B member was named in the *Who's Who Legal* list. None of the participants in Group C appeared in either *Chambers and Partners* or *Who's Who Legal's* 2014 listings of Australian competition law and economics experts. This higher prevalence of listed participants in Group A compared to Group B, and the fact that no listed experts appeared in Group C, suggested material agreement between the participant selection process and ranking scores used in this study and the identification of legal and economics experts by these two industry research publications.

At the same time, however, the methodology developed for this study assigned participants to one of three groups according to their levels of likely expertise. There was no parallel to this feature in the research methodologies used by the two industry research publications. Moreover, the ranking framework applied here extended to non-partner level professionals. This was not a feature of these other publications, but it was critical to the design of this study. Without an ability to identify both lower-level and higher-level legal specialists from a given group of individuals (and to do so in a graduated manner), a meaningful comparison between their cognitive performances would not have been possible. In these regards, neither *Who's Who Legal* nor *Chambers and Partners* could have provided sufficient data for the analysis described in the following chapters.

D *Conclusion*

This chapter described the results of the participant selection process outlined in the previous chapter. It highlighted participants' background details and explained how these were in some instances similar and in other instances different in areas relevant to

[429] The page/webpage reference for this quotation is omitted to protect this participant's anonymity.

an analysis of their cognitive abilities as competition law specialists. It also discussed the breadth and depth of this particular sample of participants, who were widely distributed geographically and who worked in different organisations, but who were also involved in providing essentially the same forms of specialist legal services at the same premium-end of the corporate law market. When initially dividing these volunteers into partners and non-partners, the main statistically significant differences between the two groups were their average years of general and specialist experience, their possession of sign-off responsibility, and the average number of merger matters in which they had been involved during their careers to date.

This information indicated that the profiles of these participants comported with the objectives of the study inasmuch as any identified differences in cognitive skills would more likely be attributable to factors other than differences in technical legal knowledge. At the same time, the areas of similarities identified provided confidence that variables other than those reflecting cognitive ability were minimized.

The results of the ranking and grouping methodologies from the previous chapter were then discussed. This included a detailed description of the manner in which individual participants were assessed and ranked against five measures of likely expertise. These rankings were then reflected in Table 4.8, which indicated which participants were in which groups according to the terminology used by Hoffman's Scheme. The highest-ranked participants in Group A (S01, S04, S05, S06, S08 and S09) were depicted as a combination of experts and masters. The mid-ranked participants in Group B (S02, S03, S07, S10, S15, S17, S18 and S19) were depicted as a combination of apprentices, journeymen, experts and masters. The lowest-ranked participants in Group C (S11, S12, S13, S14, S16 and S20) were depicted as a combination of apprentices and journeymen.

The remainder of the chapter focused on the extent to which the above groupings had preserved relevant distinctions between different levels of participants. These distinctions were confirmed to be those relating to length of professional experience, but not those relating to differences in technical legal knowledge. Comparisons were also made between the rankings generated in this study and the rankings of legal specialists published by two legal industry research organisations. These comparisons suggested that the levels of expertise amongst the higher-ranked participants were high

in an absolute sense, and that their ranking scores were in broad agreement with these industry publications notwithstanding fundamental differences in ranking methodology. At the same time, those alternative methodologies were considered incapable of categorizing all participants in this study, given the present focus on comparisons between apprentices, journeymen, experts and masters within the same specialist field of law.

The ranking of participants in this chapter, and more particularly their allocation to one of Group A, Group B or Group C, was a necessary step prior to the analysis of test results discussed in following chapters. By arranging the participants according to their relative levels of expertise, and by assuming a range of expertise from apprentice to master according to Hoffman's Scheme, the comparisons of different assessment behaviours could be related back to a participant's level of likely expertise as a competition law specialist. Identifying the categories of behaviours and performance characteristics that most clearly reveal these differences is the focus of the next two chapters.

V RESULTS

The test results presented in this chapter confirm that legal specialists with different levels of likely expertise in the same field of law do think differently when assessing legal risk in information-constrained and time-limited contexts. The observed differences include how less expert participants were more prone to laboured reasoning and incorrect assessments, and how they spent more time, on average, identifying issues than synthesising them prior to forming a legal-risk opinion. More expert participants, on the other hand, consistently provided correct assessments based on substantive legal and economic analyses, whereas the few less-expert participants who correctly assessed the level of risk in specific cases did so based on superficial analyses. Greater expertise was also associated with higher rates of verbalisation generally as well as during the identification and synthesis of issues.

Part A of this chapter provides an overview of the data collected from study participants and how these data were categorised both qualitatively and quantitatively to identify differences in how participants with different levels of likely expertise assessed legal risk. This part also explains some methodological choices that were made as part of the exploratory approach adopted during data analysis. There is also an introduction to the next two parts of the chapter which describe the qualitative and quantitative results of this analysis.

Part B records the results of the qualitative analysis concerning the apparent ease with which participants reasoned during their assessments of legal risk, the conclusiveness of their assessments, how accurate their predictions of likely outcomes were, and whether they relied on superficial or substantive legal and economic analyses when forming their views. In each of these areas, differences were identified between how participants with higher levels of likely expertise performed compared with participants with lower levels of likely expertise based on the participant rankings described in the previous chapter.

Part C details the results of the quantitative analysis that was applied to participants' think-aloud transcripts. This analysis focused on verbalisation rates and time spent on identifying issues and on drawing inferences on those issues. It covers the inter-group

comparisons between Group A and Group C participants as well as intra-group comparisons between members of these groups in the areas of ease of reasoning, certainty of assessment, assessment accuracy and depth of analysis.

Part D concludes the chapter by summarising the qualitative and quantitative differences identified. It also confirms the relevance of these differences as preliminary responses to the research question, and as contextual background to the analysis presented in Chapter 6.

A *Overview*

The 20 participants in this study generated a total of 73 think-aloud verbal transcripts containing their vocalised legal-risk assessments for test-cases A, B, C, D, E, F and G. The average time taken to complete an assessment was 7 minutes 46 seconds, which was comfortably within the 10 minute time limit loosely enforced for each test case.[430] The resulting transcripts exceeded 70,000 words in total, with an average of 965 words per test-case and 3,522 total words per participant,[431] each of whom completed, on average, 3.65 test cases.[432]

To ensure that the analysis of these data involved comparisons between legal specialists with different levels of likely expertise, the principal focus was on differences in assessment behaviour between Group A and Group C participants, as Group B participants were considered a likely mix of various levels of specialist expertise as noted in the previous chapter. This focus facilitated the search for associations between the traditional levels of expertise development described in Hoffman's Scheme on the

[430] The least time taken to provide an opinion on a case was 1 minute and 38 seconds (Participant S01 for Case F). The most time was 17 minutes and 8 seconds (Participant S19 for Case D).

[431] Generally only whole words were transcribed. Participants' vocalizations such as 'oh,' 'um,' 'mmm' and time-fillers such as 'dut dut dut' uttered while scanning a page, were only occasionally transcribed. Because the original audio recording could always be replayed by clicking on the relevant text in a transcript (a feature of the analysis software used as discussed in Section 6 of Part D of Chapter 3) these vocalizations could be audibly reviewed at any time during the analysis process and therefore phonetic transcription of non-word vocalizations did not need to be comprehensive or exhaustive.

[432] These statistics were comparable to those of Baltes et al who also required their test participants to complete more than one representative task in their think-aloud wisdom studies. In two such studies, the average time that a participant spent engaging in the set problem-solving task was 5 minutes for one study and 6.8 minutes for the other. The average number of words spoken was 561 and 699 per task, respectively. As in those studies, the verbal data collected in this study, assessed on a per-participant basis as well as in aggregate, were considered sufficient for conducting the proposed verbal protocol analyses. See PB Baltes, UM Staudinger, A Maercker and J Smith, 'People Nominated as Wise: A Comparative Study of Wisdom-Related Knowledge' (1995) 10(2) *Psychology and Aging* 155, 160; Smith, J and PB Baltes, 'Wisdom-Related Knowledge: Age/Cohort Differences in Response to Life-Planning Problems,' (1990) 26 *Developmental Psychology* 494, 499.

one hand (masters and experts in Group A and apprentices and journeymen in Group C), and, on the other hand, different cognitive skills and thinking strategies.

According to the exploratory approach used in this study, a teleological perspective was adopted to identify comparative parameters that most effectively highlighted differences between Group A and Group C participants. The aim was not to provide a comprehensive account of all possible differences, but rather to identify specific differences that were readily apparent and appeared to reflect different thinking skills and strategies relatable to differences in levels of specialist legal expertise. Both qualitative and quantitative comparisons were considered.

The qualitative measures chosen highlighted assessment behaviours that could be ascertained from reading each transcript and categorising it according to the relevant participant's ease of reasoning, the certainty of their assessments, their assessment accuracy compared with the actual outcomes of the test cases, and whether the analysis on which they relied was substantive or superficial. As this process was of a general nature and relatively straightforward to implement, it was considered unnecessary to involve a second assessor to confirm these categorisations as is usually required when individual protocols are dissected according to detailed coding procedures in word-frequency studies, for instance.[433] Instead, a descriptive approach to the discussion of individual transcripts is provided in this chapter in order to provide a more complete picture of the factors behind the categorisation decisions.

Support for the single-assessor approach is provided by Crandall, Klein and Hoffman who identify a number of circumstances in which reliance on inter-coder reliability is unnecessary.[434] Using the example of Hoffman, Coffey and Ford's analysis of weather forecasting procedures,[435] they explain that when the coding is simple or a

[433] The most recent example of this more detailed approach to protocol-analysis involving lawyers' think-aloud verbalisations, was Chay's 2006 study in which the verbal transcripts of his four participants were segmented according to list of 15 discrete codes. These codes depended on inferences about cognitive processing in the areas of control procedures, information gathering procedures, attribute identification procedures, and move identification procedures. See Allan James Chay, *Lawyer Problem Solving: An Investigation of the Knowledge Used in Solving Practical Legal Problems* (PhD Thesis, Griffith University, 2006) 92. The present study, however, focused on general performance measures that could be objectively assessed. There was also greater tolerance for errors given this study's broader research objectives and exploratory approach.

[434] Beth Crandall, Gary Klein and Robert R Hoffman, *Working Minds: A Practitioner's Guide to Cognitive Task Analysis* (The MIT Press, 2006) 102.

[435] RR Hoffman, JW Coffey and KM Ford, *A Case Study in the Research Paradigm of Human-Centered Computing: Local Expertise in Weather Forecasting* (National Technology Alliance, 2000).

comprehensive accounting of protocols is not required, a second assessor is not essential. Similarly, where the researcher is him or herself a domain expert, as was the case in the present study, confirmation of their coding decisions – particularly by a less expert third party – is unlikely to add much further value either methodologically or in terms of data quality. It is further noted that Ericsson and Simon concede that inter-coder 'reliability does not imply validity of encoding,'[436] which is the principal issue in exploratory analyses of the kind adopted here.

In any event, the findings of this study are not intended to support conclusive statements asserting cause and effect relationships or other 'strong claims about reasoning processes,' for which higher-levels of scientific rigour are required.[437] Rather, they are intended to be observational with the aim of providing practical insights and to lay the groundwork for further research in which multiple assessors of verbal protocols will be important.

The quantitative data analysed in this study were compiled within the six weeks following the completion of test interviews. However, they were not immediately analysed as the researcher's attention switched to exploring qualitative comparisons for the next several months. Once qualitative categorisation had been settled, the quantitative data were retrieved and analysed. This is relevant for allaying concerns that qualitative categories may have relied on or been influenced by earlier quantitative assessments.

There were two types of quantitative data extracted from participants' transcribed think-aloud verbalisations. The first was their rate of verbalisation as measured by how many words per minute they vocalised. The second was the total time they spent considering a given test case as well as time spent on the tasks of identifying issues and engaging in synthesis during the process of assessing legal risk. To facilitate the analysis of both types of quantitative data, each transcript was divided into 10-second segments within which individual words were counted and relevant assessment tasks identified according to a simple, two-stage flowchart. Further details of this methodology are discussed in Part C below.

---

[436] K Anders Ericsson and Herbert Simon, *Protocol Analysis: Verbal Reports as Data* (The MIT Press, 1984/93), 205.
[437] Crandall, Klein and Hoffman , above n 434, 101-102.

B *Qualitative Results*

As noted above, legal-risk assessments by participants were categorised qualitatively in four different ways. The first was the ease with which a participant reasoned while assessing a test case. The second was how certain they were about their conclusions as to the likely outcome of the case. The third was whether those assessments that were conclusive, and which therefore contained a clear opinion, were either correct or incorrect having regard to how the particular case was ultimately decided by the relevant competition authority. Lastly, those assessments which were conclusive and correct were categorised according to whether they involved superficial or substantive legal and economic analyses.

The following discussion of the categorisation process and results includes descriptive examples to illustrate relevant assessment behaviours. Tabular summaries are also provided to give a sense of the relative incidence of particular behaviours amongst more and less expert legal specialists in Group A and Group C, respectively.

1 *Ease of Reasoning*

The analysis of participants' ease of reasoning focused on whether they engaged in laboured reasoning or unlaboured reasoning.

Laboured reasoning was characterised by extended consideration of both relevant and irrelevant issues, as well as repeated references to insufficient information. Unlaboured reasoning was, by comparison, efficient and more linear in its progression. Participant reasoning was assumed to be of this latter type by default, unless there was evidence that a participant was struggling to identify key issues or consistently referred to irrelevant information and/or repeated their earlier points of contention without resolving outstanding issues. Laboured reasoning also typically resulted in longer legal-risk assessments, although this in itself was insufficient to distinguish laboured reasoning from merely thorough reasoning.[438]

Examples of laboured reasoning included:

[438] For instance, S11 took 4.49 minutes to complete his assessment of Case B, which assessment was identified as involving laboured reasoning (as confirmed by a confused and contradictory summing up of the case), while S10 took 11.80 minutes to assess Case D yet did not engage in laboured reasoning but rather undertook a particularly thorough consideration of relevant issues.

- Participant S19's reasoning in Case D (extending over 16 minutes, or more than twice the average assessment duration) which included his admission of having 'run out of puff';[439]

- Participant S18's ten-minute assessment of the same case, half-way through which he vocalised that he was 'sort of flagging here, getting tired'[440] before commenting that 'you know, you can't, can't reach a conclusion … based on a market inquiries letter, ultimately,'[441] and then admitting three minutes later to only being able to speculate on the outcome of the case;[442]

- Participant S17's consideration of possible arguments in Case B regarding other competitors in the market (an earlier inference) followed by a discussion of: the probability that such rivals may not service the same customer base; an apparent absence 'of, I guess, aggression on behalf of the ACCC'[443] based on the wording of the market inquiries letter; and a query about whether the relevant geographic market might be broader than the ACCC suggested, which could further ameliorate concerns subject to the qualification 'although … the information provided is really pretty thin;'[444] and

- Participant S14 who began to summarise his assessment of Case D after more than 13 minutes of analysis, as follows:

> seems to me that this is something that, you know … yeah, I just don't, I just don't have a sense, to be honest, of how big these guys are compared to any other players in the market … I mean if there is another big wholesaler or a big retail chain out there … then I would think that would be enough … but if there's no other big retail chains – retailer of agriproducts – and no other wholesaler … and these guys are, you know, it's a merger of one and two … then I would think this could be a difficult merger to get through … but if there is even just one other big player out there … then instinctively I feel this is a

[439] S19D Line 355.
[440] S18D Line 138.
[441] S18D Line 157-159.
[442] S18D Line 256.
[443] S17B Line 190-191.
[444] S17B Line 201.

> doable deal … perhaps with some very small local divestitures or remedies in some local markets.[445]

Twenty risk assessments in this study were identified as involving laboured reasoning with the remaining 53 involving unlaboured reasoning. These instances of laboured reasoning represented 27% of all risk assessments recorded. The distribution of this type of reasoning, however, was skewed towards Group C participants, with 12 (50%) of that group's assessments including laboured reasoning. Group B participants demonstrated such reasoning in 8 instances (27% of the cases that group considered). None of the participants in Group A engaged in laboured reasoning.

The following table shows the incidence of both laboured and unlaboured reasoning in each of groups A, B and C.

TABLE 5.1 – Ease of Reasoning

CASES	GROUP C						GROUP B								GROUP A					
	S12	S20	S14	S13	S11	S16	S10	S18	S07	S17	S02	S15	S19	S03	S04	S01	S06	S08	S05	S09
A	≈	●	≈	≈	●	≈	●	≈	●	≈	●	●	●	●	●			●	●	●
B	≈	●	≈	●	≈	●	●		●	≈	●	●	●		●			●	●	
C	≈	●	≈	≈	●	●	●	≈	●	●	●	●	≈	●	●		●	●	●	
D	≈	●	≈	●	●	●	●	≈	●	≈	●	●	≈	●	●		●	●	●	●
E																●				
F																●				
G																●				

≈ — *Laboured Reasoning* ● — *Unlaboured Reasoning*

With one exception, all Group C participants engaged in laboured reasoning in at least one instance, with three doing so in two or more instances. The one Group C participant who did not engage in laboured reasoning (S20) was subsequently assessed as having undertaken superficial analyses in two cases, as discussed further below. By contrast, none of the assessments by Group A participants was categorised as involving either laboured reasoning or superficial analyses.

This suggested that laboured reasoning was associated with lower levels of specialist legal expertise. This has intuitive appeal inasmuch as the assessment of legal risk may be expected to be more difficult – and therefore more laboured – for less expert

[445] S14D Line 326-346.

specialists, while some may choose to pre-emptively rely on superficial analysis as in the case of S20 and other Group C participants noted in Section 4 below.

2 *Certainty of Assessment*

Participants' certainty in assessing legal-risk in the test cases was categorised as either conclusive or inconclusive, with a minority refusing to provide any assessment.

A conclusive assessment consisted of a definite statement by a participant as to the level of legal risk associated with a particular transaction. It was also accompanied by, at most, only limited qualifying statements. A conclusive assessment reflected a confident, though possibly slightly qualified, statement that a merger transaction would or would not be cleared by the relevant competition authority. From a client's perspective, such an assessment would likely be viewed as both clear and actionable.

Examples of conclusive assessments included:

- Participant S01's conclusion in Case G that 'you'd expect to get the tick in the long run … not a hard one;'[446]

- Participant S03's statement in Case C that he 'would be advising the parties in half an hour that they really don't have anything to worry about;'[447]

- Participant S08's unqualified statement in Case C that 'this is capable of clearance;'[448]

- Participant S04's assessment in Case A in which the apparent existence of low barriers to entry and broad product market definitions led him to conclude, 'it seems to me that there is a real chance of securing clearance;'[449]

- Participant S12's statement in Case B that, 'if I was acting for Toll my competition advice would be that this [transaction] is likely to be opposed by

[446] S01G Line 80-84.
[447] S03C Line 116-117.
[448] S08C Line 203.
[449] S04A Line 154-155.

the ACCC;'[450]

- Participant S17's conclusion in respect of Case D that 'this transaction has a strong argument for it to be cleared, and … the parties can have a level of confidence that it will be;'[451] and

- Participant S05's conclusion in Case B:

> I think the Commission is going to have, again, some difficulty with this merger … I think Toll is going to need to be approaching the Commission with some sort of assurance or undertaking around how they're going to conduct their business in the future … particularly in regard to their shipping service … even that may not be enough.[452]

Conclusive assessments, of which there were 28 instances representing approximately 38% of all risk assessments, were observed across all levels of likely expertise. The highest incidence was amongst Group A participants, who provided firm opinions in 12 instances (63% of all legal risk assessments contributed by this group, including S01 who considered cases E, F and G), while Group B participants contributed 9 instances (30% of all assessments by this group) and Group C participants contributed 8 instances (33% of all their assessments).

In four instances participants refused to provide any assessment of legal risk. These were recorded as neither conclusive nor inconclusive assessments. Responses in this category included:

- Participant S07's statement in respect of Case A that 'there's not enough information here, I think actually, to make an assessment of what you think the outcome would be;'[453]

[450] S12B Line 238-239.
[451] S17D Line 299-301.
[452] S05B Line 237-245.
[453] S07A Line 172-173.

- S06's conclusion in respect to Case D: 'It's impossible to advise on this matter … impossible;'[454] and

- S09 in Case A, who after identifying what he considered to be 'the key things we need to know'[455] said, 'I'm in really no position to help to give information … to give any indication of how likely they [the merger parties] are to succeed without knowing a lot more.'[456]

The following table shows the incidence of conclusive, inconclusive and 'no assessments.'

TABLE 5.2 – Certainty of Assessment

CASES	GROUP C						GROUP B								GROUP A					
	S12	S20	S14	S13	S11	S16	S10	S18	S07	S17	S02	S15	S19	S03	S04	S01	S06	S08	S05	S09
A	—	●	—	—	—	—	—	—	*na*	—	●	—	—	—	●			●	—	*na*
B	●	—	—	—	—	●	—		●	—	—	—	●		—			●	●	
C	—	●	—	—	●	●	—	●	●	●	●	—	—	●	●		●	●	●	
D	—	—	—	●	●	—	—	—	—	●	—	—	—	—	—		*na*	●	—	*na*
E																●				
F																●				
G																●				

● — *Conclusive Assessment* — *Inconclusive Assessment* *na* — *No Assessment*

On a purely percentage basis, conclusive assessments of legal risk were greatest amongst more expert legal specialists and less common amongst less expert legal specialists. However, all but one Group C participant demonstrated at least the capability to provide a conclusive assessment of legal risk, with half of these participants (S20, S11 and S16) doing so in 50% of the cases they considered. While S01 and S08 in Group A also did so in 100% of the cases they considered, none of the other Group A participants was better than those three Group C participants on this measure.

Of course, a conclusive yet inaccurate or incorrect assessment cannot be equated with an accurate or correct one. Even a complete novice could provide a conclusive but

[454] S06D Line 76-78.
[455] S09A Line 163.
[456] S09A Line 167-168.

wrong answer to the problems posed. The issue of assessment accuracy is discussed in the next section.

Those participants who refused to provide an assessment of legal risk in this study were either in Group A or Group B. No Group C participant refused to complete this part of their task. While all participants were specifically requested and reminded of the ultimate aim of their assessment of the test cases, participants S06, S07 and S09 considered the task impossible to complete based on the insufficient information available and their own lack of familiarity with the relevant industries.

One explanation for this was the relative seniority of these three participants, all of whom had more than 30 years professional experience and were amongst the study's oldest participants. These individuals may therefore have been less likely to comply with a request to perform the required task than their less senior and younger colleagues, all of whom at least attempted some form of either conclusive or inconclusive assessment. Complaints regarding the difficulty of the task in the absence of more information were common across all participants, but only those most self-assured – or perhaps more predisposed to intransigence – expressly refused to provide an assessment.

3 *Assessment Accuracy (Conclusive Assessments Only)*

Where a participant provided a conclusive assessment of legal risk in a test case, it was possible to determine the accuracy of that assessment by comparing it to the actual outcome of the case as recorded in the official clearance statements issued by the ACCC or CC. This comparison was not possible for inconclusive assessments which were either indeterminate or so heavily qualified that these participants' views could not be ascertained objectively.

The conclusive assessments identified in this study were categorised into four different groups according to a participant's stated view or opinion on the likely outcome or risk involved in a case. These four groups were then classified as either correct or incorrect having regard to how closely such assessments matched the known outcomes of the cases, about which none of the participants had prior knowledge given that at the time of testing neither the ACCC nor the CC had completed their own assessments.

The first two categories of opinion were those conclusive assessments which correctly predicted that the transaction in question would be cleared without any conditions or undertakings being required by the relevant competition authority. These two categories were: (a) assessments that implied or included an express statement to the effect that the test case in question was 'No problem, a straightforward case,' or (b) assessments that intimated that clearance was 'Doable, but work needed.' As all the test cases were ultimately cleared, but it was not feasible for the researcher to determine how straightforward or difficult it was for the merger parties to secure clearance from the ACCC or CC in any given instance, both these forms of assessment were considered correct assessments of legal risk.

The other two categories of legal-risk assessments were those that reflected a participant's view that: (a) 'Undertakings would likely be required,' or (b) 'The transaction is likely to be opposed.' Given that both these scenarios were inconsistent with actual outcomes of the cases, these assessments were categorised as incorrect.

The examples of conclusive assessments provided in the previous section that were categorised as correct on this basis, included participant S01's response in respect to Case G, S03 in respect to Case C, S08 in respect to Case C, S04 in respect to Case A and S17 in respect to case D. The assessments of S05 and S12 in respect to Case B, and S11 and S13 in respect to Case D were examples of assessments categorised as incorrect.

Correct assessments were, as a percentage of all assessments made by participants within a group, most prevalent within Group A, where 11 or 58% of all assessments and 10 or 91% of all conclusive assessments were of this kind. Only participant S05's assessment in respect to Case B was categorised as an incorrect conclusive assessment. Amongst Group B participants, eight or 27% of all assessments were correct in the above sense. There was one incorrect assessment made by a Group B participant (S07 in Case B). Group C participants produced four correct and four incorrect assessments, which meant that 17% of all assessments were correct and 17% were incorrect, with the remaining 66% of Group C assessments being categorised as inconclusive.

There was a further categorisation of those participants who assessed a case within the first 60 seconds of their consideration of the relevant test-case documentation. This occurred in three instances:

- Participant S16, who upon reading that no prior clearance had been sought by the acquiring party, stated at the 46-seconds mark in his assessment of Case C, 'oh, if they thought it was easy and a non-issue, then …'[457] after which he completed a superficial analysis while acknowledging, 'I'm probably heavily influenced by the fact that the parties proceeded without clearance;'[458]

- Participant S06's statement at the 27-seconds mark in his assessment of Case C, 'in my view, the chances of a court granting divestiture of any merger case in Australia – unless it is such a clear cut case – is going to be virtually zero;'[459] and

- Participant S01's statement at the 6-seconds mark in his assessment of Case F, 'oh, publishers … yes, this will be cleared.'[460]

These categorisations of correct and incorrect conclusive legal-risk assessments are shown in the following table.

TABLE 5.3 – Assessment Accuracy (Conclusive Assessments Only)

CASES	GROUP C						GROUP B								GROUP A					
	S12	S20	S14	S13	S11	S16	S10	S18	S07	S17	S02	S15	S19	S03	S04	S01	S06	S08	S05	S09
A		√									√				√			√		
B	●					●			●				√					√	●	
C		√			√	*√		√	√	√	√			√	√		*√	√	√	
D				●	●					√								√		
E																√				
F																*√				
G																√				

● – *Incorrect Assessment* √ – *Correct Assessment* *√ – *Immediate Correct Assessment*

While there was a greater percentage of correct conclusive assessments amongst Group A participants compared to Group C participants, there were four instances in which Group C participants conclusively and correctly assessed legal risk in a test case. At the same time, incorrect conclusive assessments were predominantly confined to Group C participants, although one Group A participant (S05) also provided such an

[457] S16C Line 19.
[458] S16C Line 80-81.
[459] S06C Line 17-18.
[460] S01F Line 3-5.

assessment. In general terms, however, Group A participants appeared more capable of correctly assessing the test cases than Group C participants. Given that they had no specific factual or legal knowledge advantage, this would seem to be at least in part explained by the former possessing greater legal expertise.

A further investigation was undertaken as to whether or not these correct assessments were all based on substantive analyses, or whether some were more akin to guessing. These issues are considered in the next section.

Regarding those participants who were able to formulate conclusive and correct assessments within the first 60 seconds of considering a test case (there were no examples of similarly truncated incorrect concluded assessments), there were only three instances in which this was observed, two of which involved Group A participants and one involving a Group C participant. The substantiveness of these assessments, as an aspect that could potentially distinguish the more expert of these assessments from the less expert one, is also considered below.

4 *Depth of Analysis (Correct Conclusive Assessments Only)*

Within the category of correct conclusive assessments of legal risk, some participants relied on superficial or circumstantial analyses, while others relied on substantive legal or economic analyses. This distinction was readily apparent and provided an opportunity to separate those participants who gave a correct assessment but were in effect guessing or using circumstantial observation to inform their views, from those who had engaged in reasoned and analytically compelling analyses despite the limited availability of information and time.

Examples of superficial or circumstantial analyses leading to correct conclusive assessments, included:

- Participant S20's view that because he had not heard of the merger parties in Case A, they were likely to be small businesses such that there would be 'a fairly good chance that the merger would go, will be cleared;'[461]

[461] S20A Line 103.

- Participant S07's abrupt and full acceptance of an inferred argument that the merger parties in Case C had no material overlap between their businesses despite the test-case document stating otherwise;[462] and

- The circumstantial presumption by S02, S11, S16, S17 and S20 that the fact that the merger parties in Case C had completed their transaction without prior approval from the ACCC meant that there was unlikely to be a competition problem, especially given the high-profile status of the acquirer and the likelihood that they would have sought prior legal advice from 'one of the well-respected competition teams around town,'[463] which would have advised the parties to notify 'if it was needed.'[464]

Only participants in groups B and C formed correct conclusive assessments based on superficial or circumstantial analyses. The actual number of instances where such assessments were observed, however, was not substantial. The majority were by Group C participants (four cases representing 17% of all cases considered by participants in this group), with Group B participants engaging in this form of assessment analysis in three instances, or in 10% of all legal risk assessments by these participants.

Examples of correct conclusive assessments based on substantive legal or economic analyses, included:

- Participant S08's belief that the merger described in Case C would be 'capable of clearance … on the basis that there are imports'[465] and that buying the products at issue 'has to be, I would have thought, an alternative to hiring them;'[466]

- Participant S01's reasoning in Case G that the parties were probably seeking to merge their equipment-hire businesses 'because they are struggling … and they're really looking to buy some market share in a declining market'[467] in circumstances where 'the Internet age has really, really cut into a business like

[462] S07C Line 73-76.
[463] S02C Line 58-59.
[464] S02C Line 63-64.
[465] S08C Line 191-193.
[466] S08C Line 200-201.
[467] S01G Line 59-61.

this, providing opportunities to people to pool equipment, pool resources of this kind and hire it out;'[468] and

- Participant S19's view in Case B 'that the Commission would be unlikely to have competition concerns … in this industry'[469] especially given that the likely existence other shipping service providers in the relevant market, meant 'it would be difficult to argue that the barriers to becoming a freight forwarder were sufficiently high that the Commission would'[470] oppose the proposed transaction.

Those participants who engaged in this form of legal-risk assessment were from either Group A or Group B. There were no instances where Group C participants relied on substantive analyses when forming their correct assessments of legal risk. There were five instances where a Group B participant undertook this kind of assessment. This represented 17% of all assessments undertaken by Group B participants. The 11 instances in which Group A participants formed correct conclusive assessments based on a substantive legal or economic assessment, represented 58% of all assessments made by those participants.

These statistics are reflected in the following table.

TABLE 5.4 – Depth of Analysis (Correct Conclusive Assessments Only)

CASES	GROUP C						GROUP B								GROUP A					
	S12	S20	S14	S13	S11	S16	S10	S18	S07	S17	S02	S15	S19	S03	S04	S01	S06	S08	S05	S09
A		∇									M				M			M		
B													M					M		
C		∇			∇	∇		M	∇	∇	∇			M	M		M	M	M	
D										M								M		
E																M				
F																M				
G																M				

∇ — *Superficial Analysis* **M** — *Substantive Analysis*

This table shows how correct conclusive assessments based on substantive analyses were the default or most common responses to the legal-risk assessment task by Group

[468] S01G Line 31-32.
[469] S19B Line 159-161.
[470] S19B Line 184-185.

A participants. By contrast, most Group C participants either provided inconclusive assessments or incorrect conclusive assessments. Only a minority of Group C participants provided a correct conclusive assessment, but all of these were based on superficial analyses. No Group C participant correctly assessed legal risk in any test case using analyses as substantive as those relied on by the above-indicated Group A participants.

This confirmed earlier concerns about giving equal evidentiary weighting to (a) the conclusive assessments of Group A and Group C participants, and (b) to their correct assessments of legal risk. While there were examples of conclusive assessments and correct conclusive assessments amongst both groups, only Group A participants contributed correct conclusive assessments based on substantive analyses. Moreover, they were able to do so in more than 50% of the test cases they considered, with two Group A participants (S01 and S08) producing legal-risk assessments of this type in respect to every test case they considered.

This observation suggests that in an information-constrained and time-limited context, correct conclusive legal-risk assessments based on substantive analyses are associated with (and may therefore require) higher levels of specialist legal expertise, most likely at the levels of expert and master within Hoffman's Scheme. Further, such assessments may be predicted most of the time within such a group.

The legal specialists at the apprentice and journeyman levels of expertise in this study were significantly less likely to provide correct conclusive legal-risk assessments in the same context, and only did so based on superficial analyses.

5 *Summary*

The preceding categorisation of legal-risk assessments identified a number of qualitative differences between Group A participants and Group C participants who represented legal specialists with comparatively greater expertise and legal specialists with comparatively less expertise, respectively. These differences, which related to participants' ease of reasoning, certainty of assessment, assessment accuracy (for conclusive assessments only), and depth of analysis (for correct conclusive assessments only), constituted the first responses to this study's research question, inasmuch as these results confirm that legal specialists with different levels of expertise but the same

level of technical legal knowledge can be readily identified as thinking differently from one another when assessing risk in an information-constrained and time-limited context. Moreover, these differences can be explicitly described such that they can be listed and, to an extent, measured.

A summary of these differences is presented in the Table 5.5 below.

TABLE 5.5 – Summary of Assessment Behaviours

CASES	GROUP C						GROUP B								GROUP A					
	S12	S20	S14	S13	S11	S16	S10	S18	S07	S17	S02	S15	S19	S03	S04	S01	S06	S08	S05	S09
A	≈	∇	≈	≈	•	≈	•	≈	*na*	≈	√	•	•	•	√			√	•	*na*
B	≈X	•	≈	•	≈	X	•		X	≈	•	•	√		•			√	X	
C	≈	∇	≈	≈	∇	*∇	•	≈√	∇	∇	∇	•	≈	√	√		*√	√	√	
D	≈	•	≈	X	X	•	•	≈	•	≈√	•	•	≈	•	•		*na*	√	•	*na*
E																√				
F																*√				
G																√				

√ *Correct Assessment* ≈√ *Laboured Correct Assessment* *√ *Immediate Correct Assessment* ∇ *Correct Assessment (Superficial Analysis)*
*∇ *Immediate Correct Assessment (Superficial Analysis)* X *Incorrect Assessment* ≈X *Laboured Incorrect Assessment*
≈ — *Laboured Inconclusive Assessment* • — *Unlaboured Inconclusive Assessment* *na* — *No Assessment*

This table permits some additional vertical (by participant) and horizontal (by case) observations. The apparent differences between Group A and Group C participants on an individual-by-individual basis indicate that while all but one member of the former group produced at least one substantively-based correct conclusive legal-risk assessment, none of the latter group did. Further, as individuals Group C participants were far more likely to engage in laboured reasoning, produce incorrect conclusive assessments and provide correct conclusive assessments based on superficial analyses.

In terms of differences between different test cases, neither Case A nor Case C produced any incorrect risk assessments. Case B resulted in three such assessments, including the incorrect conclusive assessment by S05 from Group A. A third of Group C participants produced incorrect conclusive assessments in respect of Case D, while the Group A participants considering this case either engaged in inconclusive but unlaboured reasoning, refused to provide an assessment or provided a correct conclusive assessment based on substantive analysis.

The highest number of correct assessments overall was for Case C, which involved a completed transaction. This case was atypical in that most clearance applications are sought before completion of the merger transaction, a point acknowledged by every participant who considered this case. This fact provided the opportunity for superficial analyses based on the inference that because the parties had not considered it necessary to seek prior approval, the ACCC would not oppose the transaction. Half of the Group C participants considering this case relied on this inference, but none of the Group A participants considering it did.

C *Quantitative Results*

Ericsson and Simon describe a direct relationship between an individual's rate of verbalisation and the stream of non-oral information to which they are attending when problem-solving.[471] This relationship permits a number of assumptions regarding cognitive function based on both the rate and the focus of an individual's verbalised protocols when assessing legal risk as in this study.

As to a participant's rate of verbalisation ('ROV'[472]), as measured by words-per-minute ('wpm'), Ericsson and Simon identify a number of studies where this metric has been used to assess cognitive performance,[473] including: Ohlsson's observation, confirming Simon and Simon's earlier findings,[474] that the ROV often remains constant for an individual's problem-solving across different problems;[475] Simon and Simon's finding that novice physicists verbalise, on average, significantly more slowly than expert physicists;[476] and studies by Sargent,[477] Montgomery and Allwood,[478] and Deffner[479] demonstrating how more difficult problems (primarily in the form of anagrams) reliably decrease verbalisation rates.

---

[471] Ericsson and Simon, above n 436, 250.
[472] According to context, the abbreviation 'ROV' is also used here for the plural 'rates of verbalisation.'
[473] Ericsson and Simon, above n 436, 251.
[474] D P Simon and H A Simon, 'Individual Differences in Solving Physics Problems,' in R Siegler (ed) *Children's Thinking: What Develops?* (Erlbaum, 1978) 325.
[475] S Ohlsson, The Cognitive Seminar: Report No 6 – Competence and Strategy in Reasoning with Common Spatial Concepts (University of Stockholm, 1980).
[476] Simon and Simon, above n 474, 325.
[477] S S Sargent, 'Thinking Processes at Various Levels of Difficulty' (1940) *Archives of Psychology* 249.
[478] H Montgomery and CM Allwood, 'On the Subjective Representation of Statistical Problems' (1978) 22 *Scandinavian Journal of Educational Research* 107.
[479] G Deffner, *Think Aloud: An Investigation of the Validity of a Data Collection Procedure* (PhD Thesis, University of Hamburg, 1983).

Johnson[480] and Durkin[481] have similarly recorded how problem-solvers reduce their ROV when cognitive activity becomes more intense, or when they are experiencing difficulties in solving a given problem. Ericsson and Simon also note how test subjects may become 'silent when … reorganising their perceptions of a problem.'[482]

There are a variety of methods for measuring wpm when assessing an individual's ROV. For this study guidance was taken from Deffner's segmentation of verbalised transcripts into 4-second intervals.[483] In this study, 10-second intervals were chosen as a more manageable unit of segmentation given the number of transcripts requiring analysis[484] and to avoid as much as possible dividing phrases between segments. In each 10-second interval the number of words, rather than the number of letters as used by Deffner, was counted.[485] This enabled the tracking of verbalisation rates and accordingly cognitive load at different stages of the legal risk-assessment process.

Regarding the counting of non-words such as 'um' and 'er,' these were not included. However, words such as 'okay' and 'yes' when used as stand-alone statements to convey an understanding of the test documentation, were counted.

As to the focus of participants during their assessment of legal risk in individual cases, the 10-second intervals used for calculating ROV presented the opportunity to categorise where a participant was focusing their attention during any given interval. The number of possible categories of focus was limited to three: (a) Reading or clarifying the provided documentation; (b) Identifying or noting specific legal or factual issues, but not in the context of forming an immediate, overall opinion; and (c) Synthesising the provided information and combining it with information drawn from long-term memory ('LTM') to form an opinion on the likely level of legal risk presented in a given test case. These categories were considered to be broadly consistent with the three stages of effective problem-solving identified by Feltovich,

[480] Johnson, ES 'An Information Processing Model of One Kind of Problem Solving' (1964) 4 *Psychological Monographs* 78.
[481] H E Durkin, 'Trial and Error, Gradual Analysis and Sudden Reorganization: An Experimental Study of Problem Solving' (1937) *Archives of Psychology* 210.
[482] Ericsson and Simon, above n 436, 252.
[483] Deffner, above n 479.
[484] With 9 hours, 45 minutes and 13 seconds of transcribed verbalisations generated by participants during this study, there were 3,403 10-second intervals to assess compared with 8,508 intervals using a 4-second interval.
[485] While not every word conformed to the 5-letter standard, it was assumed that given the level of generality to which this information would be ultimately analysed, the size of words within each 10-seconds interval was assumed likely to average-out within and across transcripts.

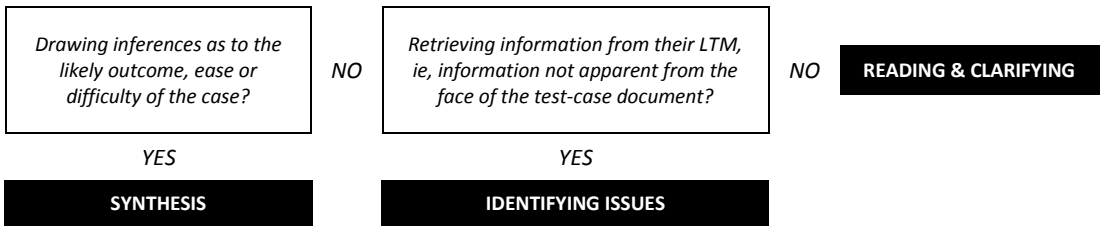
Prietula and Ericsson[486] and with Baddely's conceptualisation of the functions of working memory ('WM'), short-term memory ('STM') and LTM.[487]

With reference to these researchers' theoretical framework, a participant's reading and clarifying of the test-case documentation was assumed to involve the seeking and perceiving of data from the environment. Identifying specific legal and factual issues involved considering factors relevant to the assessment of legal risk, which necessarily involved the retrieval of relevant (and sometimes irrelevant) information from LTM before combining it with environmental data within WM. The synthesis of information perceived from the test-case documents and information retrieved from LTM in the course of forming an opinion (or indicating the further information required to form an opinion) was assumed to involve the drawing of inferences from the information available in WM (a process termed synthesis) as described in the third stage of effective problem-solving. This conceptualisation was considered theoretically sound given that such categorisation was both comprehensive and sufficient for the needs of subsequent analysis based on the above authors' problem-solving framework as described in Section (a) of Part 3 of Chapter 3 of this thesis.

The following two-stage flowchart was used to categorise each 10-second segment of each transcript as involving one of the three cognitive tasks of reading and clarifying, identifying issues or synthesis.

CHART 5.1 – Flowchart for Categorising Participant Task-Focus

Is the participant mostly …



[486] Feltovich, Paul J, Michael J Prietula and K Anders Ericsson, 'Studies of Expertise from Psychological Perspectives,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 41.
[487] A Baddeley, 'Short-Term and Working Memory,' in E Tulving and F Craik (eds), *The Oxford Handbook of Memory* (Oxford University Press, 2000) 77; A Baddeley, 'Working Memory: Theories, Models, and Controversies' (2012) 63 *Annual Review of Psychology* 1.

With respect to the classification of synthesis – the first step of the categorisation process – it was possible to confirm that this was the most likely task focus by working back from a participant's opinions on the outcome of a case. When this confirmatory process was undertaken, there was often an explicit lead-in statement made by the participant to the effect, 'So what do I think will happen in this case?' or 'So, summing all that up, I think that …' In other instances, the break between simply identifying issues – which commonly arose during the reading of the documentation – and the synthesis of the information brought into WM for the purposes of forming a view on the attendant legal risk, was usually apparent even when explicit signalling was absent.

Once a transcript had been divided into 10-seconds intervals and the words in each such segment counted (and converted to a wpm equivalent), the participant's focus was identified as being the main or primary focus for that interval using the above flowchart. This enabled both time-spent and wpm calculations to be made with respect to each task. An example of how transcripts were apportioned and categorised according to this approach is provided in Appendix C.

This approach to coding was considered a valid alternative to segmentations according to inferred mental processes[488] and attempts to identify 'ideas'.[489] Not only was it consistent with an accepted theoretical framework relating to the roles of various memory functions within the problem-solving space, it provided a simple and easily administered segmentation of protocols based solely on objectively-determined time intervals. This avoided uncertainties concerning the nature and duration of individually inferred thought processes, in which regard the study reported by Goor and Sommerfeld,[490] where the authors categorised protocols based on 3-second intervals, was both instructive and supportive of the chosen approach.[491]

With respect to the task of reading and clarifying, it was not considered helpful to rely on verbalisation rates in this area. This was because some participants read silently

[488] See Ericsson and Simon, above n 436, 204-214 for a listing of various studies of this kind.
[489] C O Smith, *The Structure of Intellect Processes Analyses Systems: A Technique for the Investigation and Qualification of Problem Solving Processes* (PhD Thesis, University of Houston, 1971).
[490] A Goor and R E Sommerfeld, 'A Comparison of Problem-Solving Processes of Creative and Non-Creative Students' (1975) 67 *Journal of Educational Psychology* 495.
[491] For instance, although these researchers segmented verbal transcripts by a strict time interval, their categorisation of these temporal protocols necessarily included contextual considerations beyond the relevant target interval. Similarly, in the present study it was useful when categorising a 10-second protocol to consider what participants were thinking in intervals before and afterwards.

with only sporadic verbalisations while others read out loud, sometimes at great speed. Comparisons between participant's ROV based on these different approaches were therefore more likely to reflect personal preferences or habits than any generalizable cognitive traits. Time spent on reading and clarifying as a proportion of overall assessment time was ultimately calculated by deducting from total time the time a participant spent identifying issues and engaging in synthesis.

The comparisons reported in this part of the chapter are of two broad types. The first involved comparing the legal-risk assessment performances of Group C participants against the performances of Group A participants. This inter-group comparison type was intended to highlight apparent differences in cognitive performance based on the different levels of expertise represented by these two groups. To use the terminology from Hoffman's Scheme, these comparisons were conceptualised as comparisons between apprentices and journeymen on one side, and experts and masters, on the other.

These comparisons, which made use of the previously identified qualitative differences in legal-risk assessment approaches, were as follows: (a) All Group C participants against an all Group A participants baseline; (b) Laboured reasoning by Group C participants against an all Group A participants baseline; (c) Unlaboured reasoning by Group C participants against an all Group A participants baseline; (d) Unlaboured reasoning, excluding superficial analyses, by Group C participants against an all Group A baseline; (e) Superficial analyses by Group C participants against an all Group A baseline; and (f) Incorrect assessments by Group C participants against an all Group A baseline.

While it was possible to undertake further comparisons involving conclusive assessments, inconclusive assessments, correct assessments and incorrect assessments, this was considered likely to significantly expand the scope of the thesis and lead to a proliferation of statistics of questionable marginal value. As the aim was to highlight a more manageable number of differences between how legal specialists with different levels of likely expertise think differently from one another, only the above

comparisons of an inter-group type were undertaken. These comparisons were considered the most relevant based on the prior inter-group results.[492]

The second broad type of comparison documented in this part of the chapter involved intra-group comparisons. First was a comparison between the best Group A participants against the rest of the Group A participants as a baseline. For the purposes of this comparison, those Group A participants who had formed correct assessments of legal risk based on substantive analyses were considered more expert than those Group A participants who had provided an inconclusive or incorrect assessment, or who did not provide any assessment. While this may be considered a simplistic distinction, the objective was to identify what those Group A participants who achieved the best results (in terms of accuracy, reasoning methodology and usefulness for their hypothetical clients) did differently from their peers. In effect, it was an attempt to determine if the cognitive traits that distinguish masters from experts could be hypothesised and explained using statistical data.

A further level of intra-group comparison was between Group C participants, who were all assumed to be within the apprentice and journeyman categories according to their relative levels of specialist legal expertise. The objective here was to understand better Group C participants' propensity for laboured assessments, which as discussed in the previous part of this chapter was an obvious distinguishing feature of Group C assessment behaviour compared with Group A participants' assessment approaches. By comparing those Group C assessments that involved laboured reasoning against those that did not, the aim was to identify those cognitive differences indicative of this kind of reasoning within a group of similarly-ranked legal specialists.

Lastly, the following sections include a considerable amount of quantitative detail. Rather than consign this information to an appendix, it was considered important to include it here for two reasons. First, it furthers the open analytical approach of this thesis and avoids obscuring the underlying methodological decisions on which its findings rely. Second, the graphical information presented below is central to the analysis undertaken in subsequent chapters, which confirmed the value of presenting its

[492] In fact, comparisons across almost every possible inter-group and intra-group combination were undertaken as part of the background research to this study. However, they were generally not reported to keep the thesis within the specified word-limit. Where relevant and helpful in explaining particular phenomena, incidental references to data from these comparisons are noted in subsequent chapters.

associated quantitative data here – where it serves a necessary explanatory function – rather than including them simply as addendum entries.

1 *Inter-Group Comparisons between Group C and Group A participants*

(a) All Group C Against Group A Baseline

A number of statistically significant differences were identified when comparing the ROV and task focus of all Group C participants against the baseline metrics of participants in Group A. Other differences were observed, but these were not found to be statistically significant and therefore the null hypothesis that there was no difference could not be dismissed.[493] The following statistics are reflected in Graph 5.1 further below, and references to columns are to the numbered columns in that graph.

In terms of verbalisation rates, as measured by wpm, Group C participants as a whole and across all test cases were 12% slower on average (Column 1). Whereas Group A participants averaged 126 wpm, Group C participants averaged only 111 wpm. Group C participants were also 17% slower in their verbalisation rates, on average, when identifying issues (118 wpm compared to Group A's 142 wpm) (Column 2) and 8% slower when engaging in synthesis (129 wpm compared to 140 wpm)(Column 3).

There were no statistically significant differences in average ROV between Group A and Group C participants as regards identifying issues as a ratio of their average ROV for cases overall (Column 4). In other words, both groups appeared to verbalise more quickly when identifying issues compared to their average ROV (at around 110% of their overall ROV), and there was no statistically significant difference as to this actual increase.

Similarly, both groups' ROV when identifying issues as a ratio of their average synthesising ROV (Column 5) could not be statistically differentiated (again, both groups verbalised faster when synthesising compared to their overall ROV).

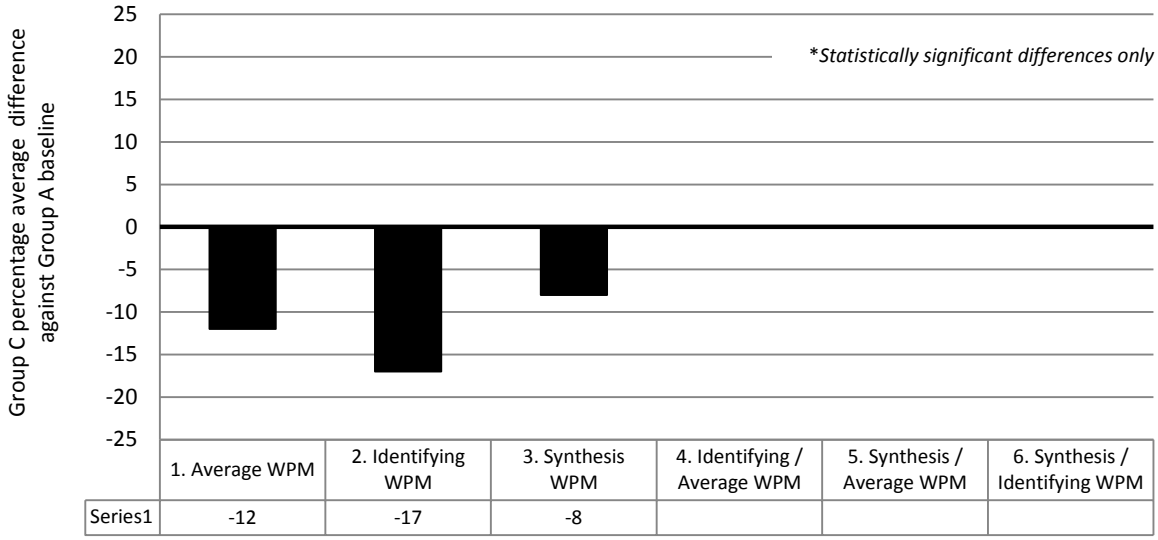
There was also no statistically significant difference identified between the ROV of Group C and Group A participants in terms of their ratio of synthesis wpm to average

[493] As previously stated in Part A of Chapter 4 (see above n 383), all tests for statistical significance in this study took the form of Student T-Tests with a null hypothesis of 'no difference between the two sample groups' and assumptions of a two-tailed distribution, two-sample unequal variance, and normal distribution. The confidence interval was set at $\alpha = 0.05$ (95% confidence level).

wpm (Column 6). Even though it appeared that the ROV of Group A participants was on average the same regardless of whether they were identifying issues or engaging in synthesis and that Group C participants appeared to verbalise 12% faster when synthesising than when identifying issues, there was no statistically significant difference between the groups on this metric.

These results are presented graphically as follows:

GRAPH 5.1 – All Group C Against Group A Baseline: Rate of Verbalisation*



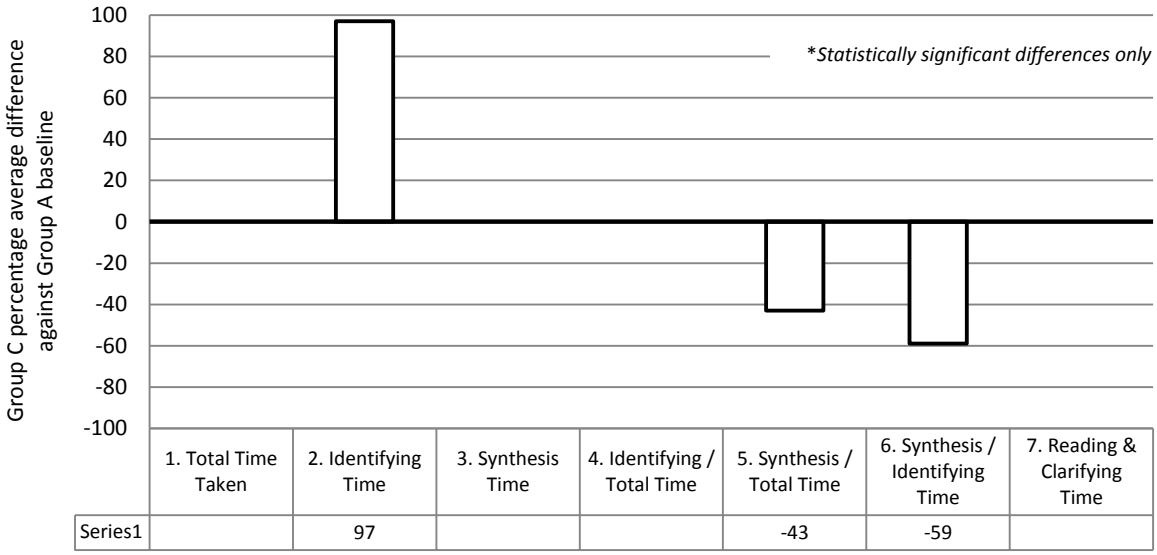
Graph 5.1 is constructed around the ROV of Group A participants as the baseline for comparisons with Group C participants in each of the above areas identified in columns 1 to 6. Where Group C participants as a whole and on average were observed to have a lower ROV than Group A participants, the columns extend downwards to the extent of the percentage difference. As can be seen from this graph, Group C participants had lower ROV on the first three measures relating to: average or overall wpm when attending to a test case (Column 1); average wpm when identifying issues (Column 2); and average wpm when synthesising issues (Column 3).

Columns 4, 5 and 6 of this graph record when participants in either Group C or Group A were observed verbalising appreciably faster or slower when identifying issues or engaging in synthesis compared to their overall ROV or when engaging in the other task. Where such differences were statistically significant (or not significant as in the

above case), inferences may be drawn concerning the cognitive performance or strategies employed by participants with different levels of likely expertise.

The next graph shows how, on average, Group C participants allocated their time to reading and clarifying the test-case documents, identifying issues and synthesising issues in a manner different from Group A participants, who were assumed to possess greater specialist legal expertise.

GRAPH 5.2 – All Group C Against Group A Baseline: Time Allocation*



Graph 5.2 records how Group C participants spent, on average across all test cases, almost double the amount of time identifying issues compared to the time that Group A participants spent on this task. Specifically, in Colum 2 above, Group C participants are shown to have spent on average 97% more time on this task (2.74 minutes vs 1.39 minutes). This is notwithstanding there was no statistically significant difference between the average total-time taken by Group C and Group A participants in completing their assessments.

This difference in identifying time was one of three areas where statistically significant differences were identified between Group C and Group A participants in relation to how they allocated their time when assessing the test cases.

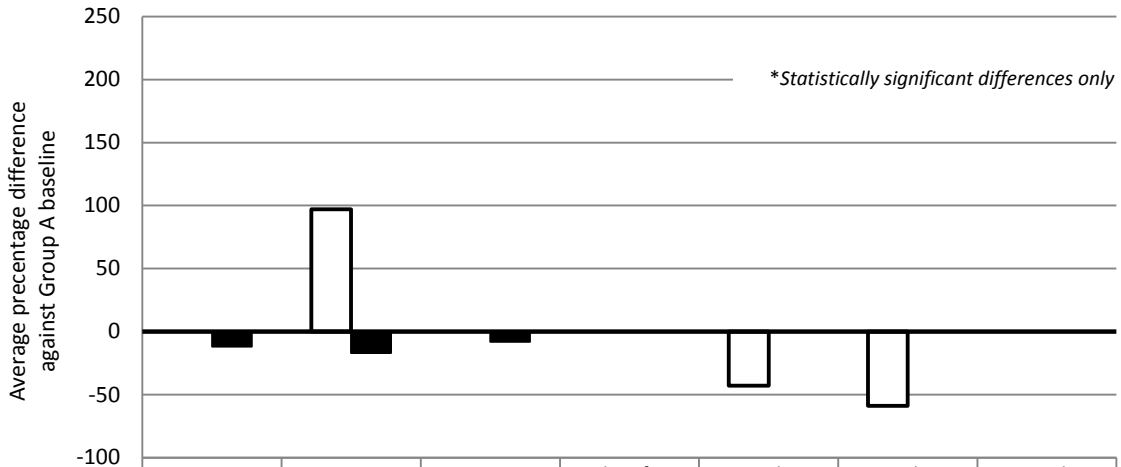
As a ratio of total time, Group C participants also spent 43% less time on synthesis than Group A participants (24% vs 42% of total assessment time)(Column 5). In respect to

synthesis time compared to time spent identifying issues, their ratio (of 138%) was 59% smaller than Group A participants' ratio (of 340%)(Column 6). In other words, while Group C participants spent on average almost 1.4 more time on synthesising issues than identifying issues, Group A participants spent on average 3.4 times more time on synthesis than on identifying issues.

The overall picture presented by these results is that Group C participants when compared with Group A participants, were more focused on – and allocated significantly more time to – identifying issues than on synthesising them when assessing legal risk in the context of this study.

Having explained the form and significance of the above graphs, for greater efficiency these visualisations can be combined to form the following multi-purpose graph.

GRAPH 5.3 – All Group C Against Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	0	97	0	0	-43	-59	0
WPM % Diff	-12	-17	-8	0	0	0	0

This graph permits the further observation that although Group C participants were, on average, slower than Group A participants when identifying issues in terms of their ROV, they spent proportionally longer considering such issues (Column 2).

As with the previous two graphs, the interpretation of Graph 5.3 requires an appreciation for the nature of percentage comparisons. Whereas in Column 2 it is clear that Group C participants spent, on average, almost double the amount of time identifying issues compared to Group A's average time, it is less obvious that a

differential of comparable magnitude exists in Columns 5 and 6 in relation to the comparison of synthesis ratios. Because these two other columns measure negative percentages, they can never be more than 100. If the baseline was to be reversed, such that Group C participants' performances were the benchmark, the resulting positive reading would be that Group A participants spent on average around 100% more of their time, in proportional terms, on synthesis compared to overall time spent – or double that of Group C participants.

Accordingly, a mental conversion is required to ensure that negative readings below the baseline are not discounted because of presentational limits inherent in this graphical format.

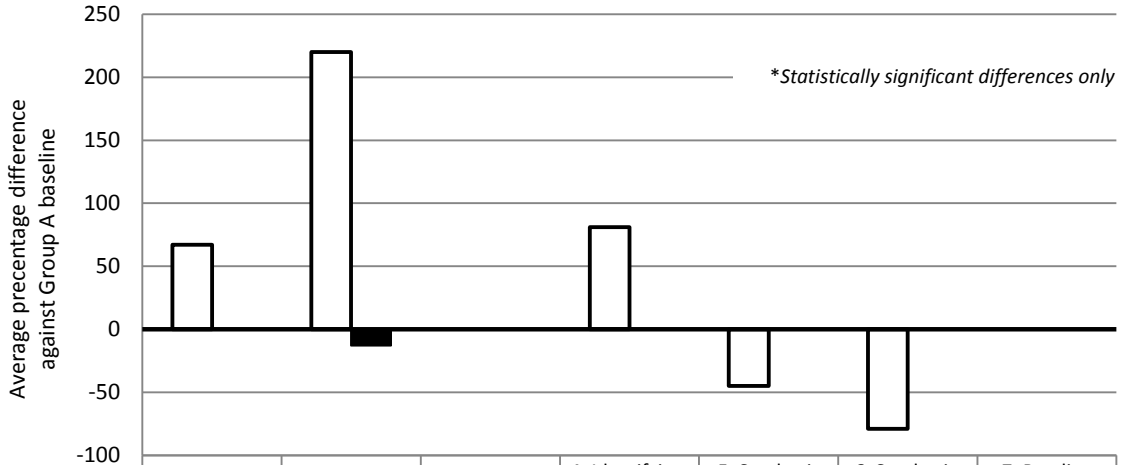
This combined graph presentation will be used to record the quantitative analysis of time and task data in the remainder of this chapter.

(b) Group C Laboured Reasoning Against Group A Baseline

Laboured reasoning was identified in Part B of this chapter as a behaviour that most Group C participants engaged in, but which no Group A participants did. It was therefore considered a defining characteristic of how less expert legal specialists assessed legal risk in this study compared to how more expert specialists did.

The following graph shows the results of the quantitative comparison between those instances in which Group C participants engaged in laboured reasoning and the unlaboured reasoning of all Group A participants.

GRAPH 5.4 – Group C Laboured Reasoning Against Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	67	220	0	81	-45	-79	0
WPM % Diff	0	-13	0	0	0	0	0

Unsurprisingly, Group C participants who engaged in laboured reasoning took, on average, 67% longer overall to assess legal risk in the test cases when compared against Group A participants (11.31 minutes vs 6.76 minutes)(Column 1). However, it was not apparent that this was because they had a lower overall ROV, as no statistically significant difference in this metric was found in terms of their overall assessment of test cases. Rather, it would appear that they took longer to complete their assessments because they verbalised more, not slower.

This difference in time spent assessing cases was ostensibly the result of these Group C participants spending, on average, over three times as many minutes identifying issues than Group A participants (4.45 minutes vs 1.39 minutes)(Column 2). It can also be inferred that they worked harder when identifying issues inasmuch as their ROV was 13% less, on average, compared with Group A participants on this task (124 wpm vs 142 wpm)(Column 2). However, this percentage difference cannot explain the total observed difference in time taken on this task.

Group C participants who engaged in laboured reasoning spent essentially the same amount of time – in absolute terms – engaged in synthesis as Group A participants, at least insofar as no statistically significant difference was found on this measure (Column 3). But this appears to be because they spent more time assessing cases overall. As a proportion of total time spent, these Group C participants allocated 45%

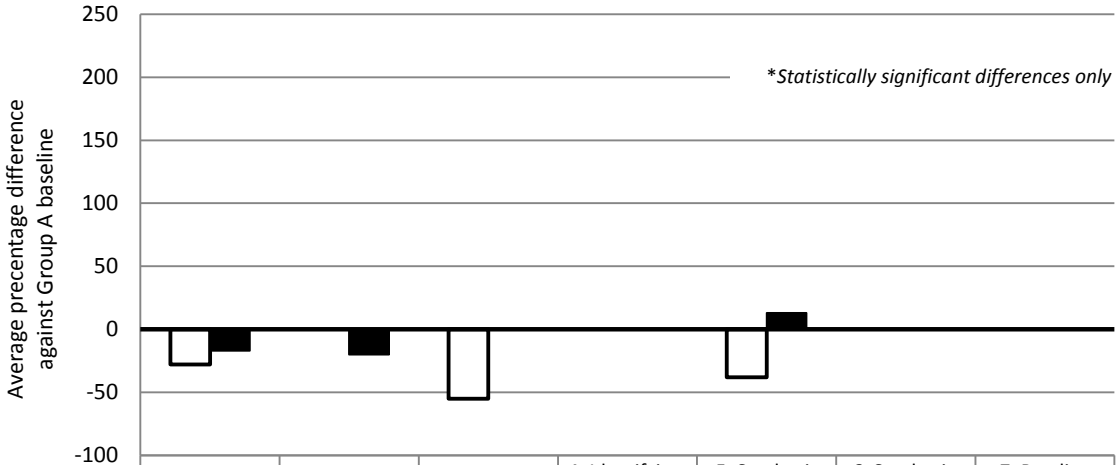
less time, on average, to synthesising issues compared to Group A participants (23% vs 42%) (Column 5). Given that no meaningful difference was detected in relation to their reading and clarifying time compared to Group A participants (Column 7), this reduction in synthesis time was wholly or largely the result of their having allocated over 80% more of their overall assessment time to identifying issues (38% vs 21%)(Column 4).

This conclusion was confirmed by these Group C participants spending only two thirds of the time they spent on synthesis on identifying issues compared to Group A participants, who spent 3.4 times more time on synthesis than on identifying issues (70% vs 340%)(Column 6).

(c) Group C Unlaboured Reasoning Against Group A Baseline

Those Group C participants who engaged in unlaboured reasoning did not spend any more or any less time than Group A participants identifying issues, although their ROV was 20% slower when undertaking this task (113 wpm vs 142 wpm)(Column 2 in Graph 5.5 below). Their slower ROV in this context would seem to explain their overall slower ROV when assessing legal risk in the test cases (105 wpm vs 126 wpm)(Column 1) given that their ROV during synthesis was, on average, indistinguishable from that of Group A participants' (Column 3) and their ROV was higher when synthesising compared to their overall ROV (Column 5). In other words, they sped-up more when synthesising compared to their overall ROV, but this was because their overall ROV was lower than that of Group A participants. As noted, there was no statistically significant evidence that they actually synthesised at a higher ROV relative to participants in Group A (Column 3).

GRAPH 5.5 – Group C Unlaboured Reasoning Against Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	-28	0	-55	0	-38	0	0
WPM % Diff	-17	-20	0	0	13	0	0

However, this subgroup of Group C participants took 28% less time, on average, than Group A participants to complete their assessments of legal risk (4.85 minutes vs 6.76 minutes)(Column 1). This was despite their overall ROV being 17% lower. The specific area in which they appeared to cut back was synthesis. With respect to this task, unlaboured assessments by Group C participants involved devoting 55% less time to synthesis, in absolute terms and on average, than Group A participants (1.24 minutes vs 2.76 minutes)(Column 3) and 38% less time as a proportion of their overall assessment time (26% vs 42%)(Column 5). At the same time, the fact that their ROV was 13% higher than that of Group A participants when synthesising compared to their overall ROV does not fully explain this reduced time spent on synthesis (Column 5).

When compared to Group A participants' legal-risk assessments, assessments by Group C participants who engaged in unlaboured reasoning appeared to differ from those who undertook laboured assessments mainly in their reduced allocation of time for identifying issues and for completing their overall assessments. This overall reduced time came mainly at the expense of time spent on synthesis. It may therefore be hypothesised that the unlaboured nature of these participants' assessments was related to their doing less in terms of identifying issues and in doing less (and faster) synthesis.

(d) Group C Unlaboured Reasoning (excluding superficial analyses) Against Group A Baseline

After excluding superficial analyses from those assessments identified as involving unlaboured reasoning, there was a further reduction of differences between unlaboured reasoning assessments and the assessments by Group A participants. This suggested that unlaboured reasoning involving superficial analyses, of which there were only four instances amongst Group C participants, involved a different risk assessment approach in terms of the cognitive measures applied here.

The absence of these four assessments from the comparative analysis between unlaboured assessments and Group A assessments eliminated the previously noted overall difference in time taken to complete the test cases between these two groups – or at least a statistically significant difference. This suggested that the previous observation that Group C unlaboured assessments were, on average, 28% shorter in duration than Group A assessments, was mostly if not wholly attributable to the superficial analyses which were a subset of unlaboured assessments.

(Superficial analyses are separately compared against Group A assessments in the next section.)

The following graph provides a more complete picture of unlaboured assessments with superficial analyses excluded, compared with Group A assessments.

GRAPH 5.6 – Group C Unlaboured Reasoning (Excluding Superficial Analyses) Against Group A Baseline*

	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	0	0	-42	0	0	0	0
WPM % Diff	-16	-22	-9	0	0	0	0

As noted in the comparison between all Group C assessments and all Group A assessments, Group C participants as a whole had lower ROV than Group A participants. This was not the case, however, in relation to laboured assessments, which was somewhat unexpected. Participants engaging in unlaboured assessments, on the other hand, whether including or excluding superficial assessments, verbalised on average 16% slower overall (106 wpm vs 126 wpm) and 22% slower when identifying issues (111 wpm vs 142 wpm) as compared with Group A participants. With the exclusion of assessments involving superficial analyses, the ROV of unlaboured assessments was 9% lower, on average, compared to Group A assessments (128 wpm vs 140 wpm)(Column 3 above).

The other apparent constant was less time spent on synthesis. Unlaboured assessments inclusive of superficial analyses involved 55% less time on synthesis compared to Group A assessments (see Column 3 of Graph 5.5). Unlaboured assessments excluding superficial analyses, on the other hand, involved 42% less time (1.54 minutes vs 2.76 minutes)(Column 3). This reflected the considerably shorter duration of superficial analyses.

However, there was no statistically significant difference on this measure when considering either all Group C participants or simply those who engaged in laboured assessments (refer sections (a) and (b) above). Presumably this was because those

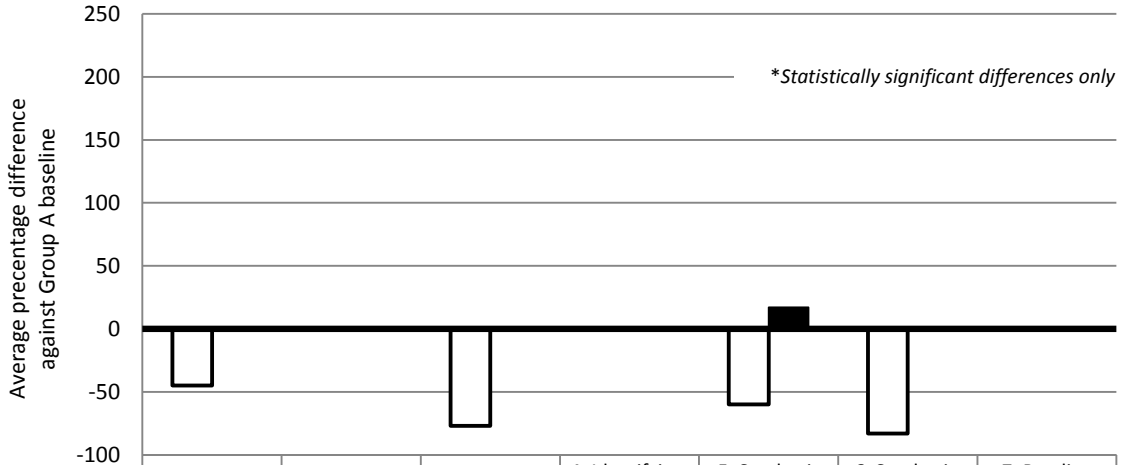
assessments were longer overall and therefore in absolute terms these participants spent longer on synthesis, even though in relative terms they spent less time as a proportion of total time and time spent on identifying issues as observed in the previous sections.

Because unlaboured assessments were generally shorter than other Group C assessments, absolute allocations of time spent on various tasks were less. However, in proportional terms, as noted in Columns 4 to 7 in Graph 5.6 above, unlike the other Group C assessments there was no significant difference between unlaboured assessments (excluding superficial analyses) and Group A assessments on this measure. This again raised questions about the differences between Group C assessments involving superficial analyses and Group A assessments. This comparison is discussed next.

(e) Group C Superficial Analyses Against Group A Baseline

Group C assessments involving superficial analyses were small in number (as previously noted, there were only four such assessments), although statistically they had a disproportionate significance in the context of the present analysis. Mindful of both these factors (a small sample size with a big statistical effect), comparisons were undertaken between these assessments and the Group A baseline, the results of which are summarised in the following Graph 5.7.

GRAPH 5.7 – Group C Superficial Analyses Against Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	-45	0	-77	0	-60	-83	0
WPM % Diff	0	0	0	0	17	0	0

Unlike the other categories of Group C assessments, participants engaging in superficial analyses were not significantly slower than Group A participants in terms of their overall verbalisation rates nor in terms of their ROV when identifying issues or engaging in synthesis. Moreover, they appeared to be the only category of Group C participants assessed to this point who were appreciably faster – 17% faster, on average – than Group A participants when synthesising compared to their overall ROV (129% vs 110%)(Column 5). In other words, they sped up more than any of the other Group C or Group A assessment types relative to their overall ROV when engaging in synthesis. This was presumably because their synthesis was more superficial and therefore cognitively less demanding.

Notwithstanding that their ROV were not significantly higher in relation to any other measures (relative to Group A participants), those Group C participants who engaged in superficial analyses spent substantially less time than Group A participants engaging in synthesis in both absolute and relative terms. Again, this was ostensibly a hallmark of their shorter and more superficial approaches.

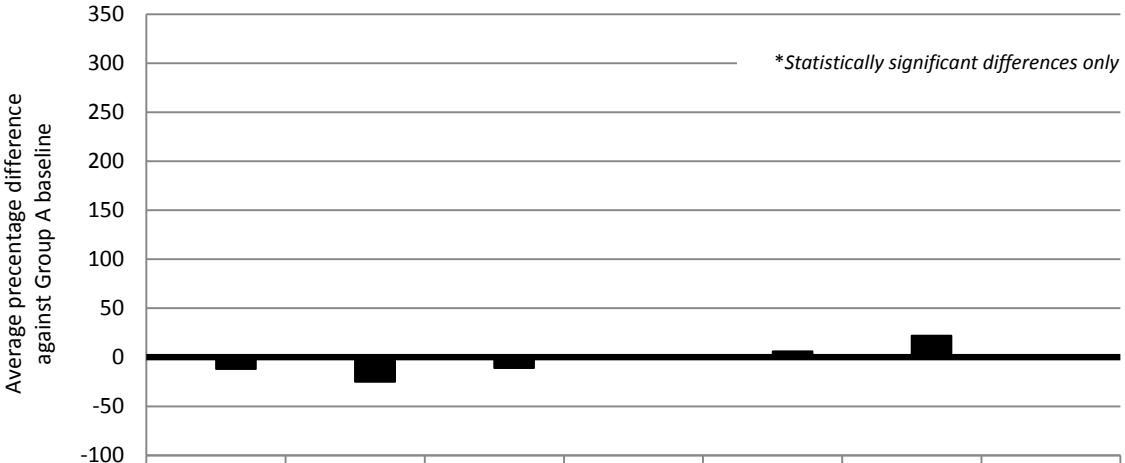
Overall, these Group C participants spent, on average, 45% less time considering cases compared to Group A participants (3.72 minutes vs 6.76 minutes)(Column 1) and 77% less time engaging in synthesis (0.63 minutes vs 2.76 minutes)(Column 3). They also spent 60% less time synthesising as a proportion of overall time (17% vs 42%)(Column 5), which was only partly explained by their higher ROV in this area. Arguably the most significant difference, however, was their spending 80% less time synthesising relative to their time spent identifying issues when compared to Group A participants (59% vs 340%)(Column 6). In this regard, they appeared to be much more strongly biased towards identifying issues than on synthesising them compared to their more expert colleagues.

(f) Group C Incorrect Assessments Against Group A Baseline

Incorrect assessments were small in number relative to all assessments, with four such assessments provided by Group C participants and one by a Group A participants. The interpretation of comparative data based on this sample must therefore be undertaken carefully, as was the case with assessments involving superficial analyses. The following graph shows the average percentage differences between the performances of

the Group C participants in these incorrect assessments as measured against all Group A assessments.

GRAPH 5.8 – Group C Incorrect Assessments Against Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	0	0	0	0	0	0	0
WPM % Diff	-12	-25	-11	0	6	22	0

One obvious feature of these comparative data is the absence of statistically significant differences in terms of time spent considering the relevant test cases overall or identifying issues or the synthesis of those issues. In these areas, the four incorrect assessments by Group C participants are statistically indistinguishable on the above cognitive measures.

But the apparent ROV differences are arguably just as significant, if not more so. While the sample size is small, as shown in Column 2 the 25% lower ROV amongst these participants when identifying issues in the course of making incorrect assessments was unusual. None of the preceding Group C comparisons against the same baseline yielded this large an average difference. These incorrect assessments were associated with an average ROV for identifying issues of 107 wpm, compared to an average of 111 wpm for unlaboured reasoning excluding superficial analyses, which was the next lowest ROV for this task. The lower average ROV overall when assessing cases of 111 wpm (compared to Group A's 126 wpm) and when synthesising issues of 129 wpm (Group A 140 wpm) are similar to the results for the other Group C comparisons above.

The 6% higher ROV when synthesising compared to overall average ROV when assessing these test cases (Column 5) is not a significant difference compared to other Group C comparisons against the Group A baseline. However, the 22% higher average ROV when synthesising compared to identifying issues is significant (122% vs Group A 100%)(Column 6). Ostensibly, this is more reflective of their lower ROV when identifying issues (Column 2) than their higher ROV when synthesising.

While not statistically significant, but in the interests of completeness, the following statistics relate to the sole incorrect assessment by a Group A participant.

Participant S05 incorrectly assessed high legal risk in Case B. In doing so, his overall ROV was 141 wpm (which was significantly higher than the average overall ROV of 111 wpm for the Group C incorrect assessments[494]). In addition: S05's average ROV when identifying issues was 149 wpm (compared to 107 wpm for the Group C incorrect assessments[495]); his average ROV when synthesising was 154 (compared to 129 wpm[496]); and, his ROV when synthesising compared to his ROV when identifying issues was 103%, which was closer to the Group A average of 100% than to the average amongst Group C incorrect assessments of 122%.

Apart from the fact that he incorrectly assessed the level of legal risk in Case B, in terms of the above ROV measures S05 otherwise appeared to perform more like his Group A colleagues than the Group C participants who also provided incorrect assessments.

(g) Summary

The first comparison in this part was between all Group C participants and all Group A participants. The quantitative differences identified and shown in Graph 5.3 related first to Group C participants' lower average ROV when assessing the test cases overall, when identifying issues, and when engaging in synthesis. In relation to time spent on these tasks, Group C participants spent, on average, almost double the time identifying issues compared to Group A participants, but substantially less time engaged in synthesis as compared to their overall average time spent assessing cases and as

[494] The highest overall ROV amongst Group C incorrect assessments was 118 wpm for S11D.
[495] The highest ROV when identifying issues amongst Group C incorrect assessments was 119 wpm for S12B.
[496] The highest ROV when synthesising amongst Group C incorrect assessments was 134 wpm for S11D.

compared to their time spent identifying issues. These were the generalised quantitative differences identified between, on the one hand, the assumed apprentices and journeymen in Group C and, on the other, the experts and masters in Group A.

When isolating the performances of those Group C participants who engaged in laboured reasoning, similar and more exaggerated results were recorded. In terms of ROV, however, laboured reasoning was associated with higher verbalisation rates than for Group C participants generally. Specifically, there was no statistically significant difference between Group C laboured assessments and the assessments by Group A participants in terms of overall ROV, nor in their ROV during synthesis.[497] This suggested that laboured reasoning, as it had been defined and applied in the categorisation of legal-risk assessments in this study, did not necessarily entail greater cognitive intensity or cognitive load. Instead, it was associated with spending three times longer identifying issues[498] (explainable in part because laboured reasoning typically, though not always, resulted in significantly longer overall assessment times), and more than double the amount of time spent on this task compared to (i) time spent overall, and (ii) time spent on synthesis.[499]

Unlaboured assessments by Group C participants resulted in other statistically significant differences against the Group A baseline. Unlaboured reasoning involved lower ROV, on average, in terms of the overall assessment and in terms of identifying issues.[500] However, when compared to Group A participants, unlaboured reasoning was found to involve a more significant increase in ROV when synthesising compared to average ROV for the assessment process overall.[501] In this respect, unlaboured reasoning appeared to be associated with less cognitively intensive synthesis and more cognitively intensive issue identification. At the same time, these Group C participants spent less time assessing cases overall, less time engaging in synthesis and less time engaging in synthesis as a proportion of total assessment time when benchmarked against the Group A participants' baseline performances in these areas.[502]

[497] Columns 1 and 3 of Graph 5.4.
[498] Column 2 of Graph 5.4.
[499] Columns 5 and 6 of Graph 5.4.
[500] Column 1 of Graph 5.5.
[501] Column 5 of Graph 5.5.
[502] Columns 1, 3 and 5 of Graph 5.5.

Lastly, quantitative comparisons were undertaken between legal-risk assessments involving unlaboured reasoning excluding superficial analyses, superficial analyses alone, and incorrect assessments against the Group A baseline. These comparisons yielded further confirmation that unlaboured reasoning in a substantive sense was associated with lower ROV overall, when identifying issues and when engaging in synthesis.[503] It was also associated with less time spent on synthesis.[504] Otherwise in terms of the allocation of time across the tasks of identifying issues and reading and clarifying information, assessments involving unlaboured reasoning were, on average, statistically indistinguishable from Group A assessments as a whole.

Superficial analyses were associated with much shorter assessments compared to the baseline assessments, and with much less time spent on synthesis in absolute terms, as a proportion of overall assessment time and as compared to time spent identifying issues.[505] Superficial analyses otherwise involved statistically the same ROV as the baseline assessments, except for a comparatively higher ROV when synthesising compared to overall ROV.[506] Incorrect assessments by Group C participants involved significantly lower ROV when identifying issues compared to the baseline statistics and to other Group C assessments.[507] However, it was noted that the samples sizes for both superficial analyses and incorrect assessments were small in statistical terms.

2 *Intra-Group Comparisons*

The focus of the preceding discussion were comparisons between legal-risk assessment behaviours associated with different levels of specialist legal expertise as represented in groups A and C. These inter-group comparisons revealed quantitative differences in several of the areas where qualitative differences between Group A and Group C participants had been identified. Where these differences were found to be statistically significant, hypotheses were postulated concerning the role played by participants' different levels of expertise in creating these differences.

In this section, a different approach is adopted. The focus here is not on gross expertise differentials, but on differences between different legal-risk assessment approaches and

[503] Columns 1, 2 and 3 of Graph 5.6.
[504] Column 3 of Graph 5.6.
[505] Columns 1, 3, 5 and 6 of Graph 5.7.
[506] Column 5 of Graph 5.7.
[507] Column 2 of Graph 5.8.

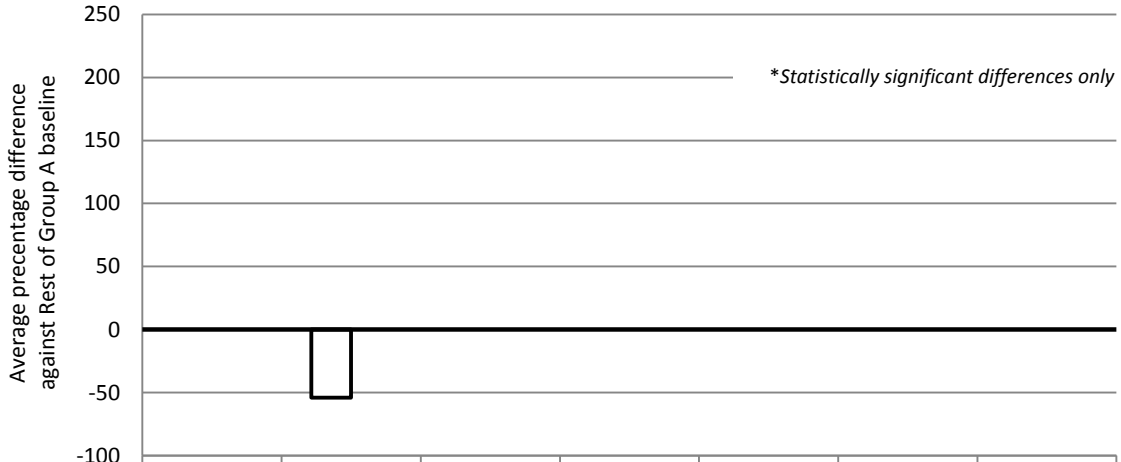
strategies amongst participants in the same group. Having identified, for instance, the higher incidence of laboured reasoning amongst lower ranked legal specialists (Group C participants) compared with higher ranked ones (Group A participants), the nature of laboured reasoning was investigated further by comparing the quantitative performances of participants who differed on this measure, but who otherwise possessed a similar level of specialist legal expertise by virtue of being allocated to the same group of similarly ranked participants.

The first comparison of this kind sought to identify what distinguished the best Group A assessments from the other assessments by Group A participants. This is followed by comparisons concerning the nature of laboured, unlaboured and superficial assessments amongst Group C participants.

(a) Best of Group A Against Rest of Group A

There was only one statistically significant difference found between those Group A participants who provided accurate legal-risk assessments based on substantive analyses (the so-called Best of Group A) and the rest of Group A participants' assessments.[508] This difference concerned the amount of time spent identifying issues, as shown in the following graph.

GRAPH 5.9 – Best of Group A Against Rest of Group A Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	0	-54	0	0	0	0	0
WPM % Diff	0	0	0	0	0	0	0

[508] The number of assessments in each category was 11 for the Best of Group A and 7 for the Rest of Group A.

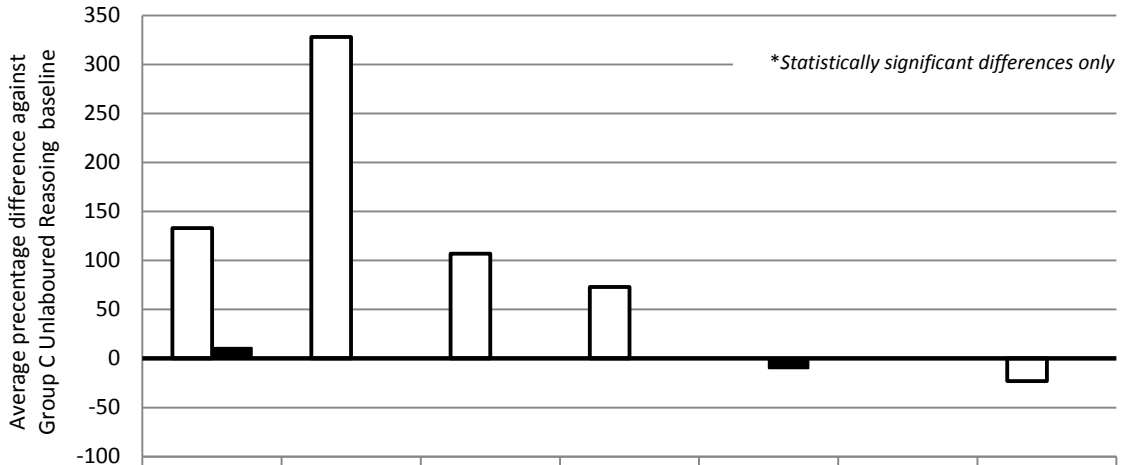
Whereas in all other respects both categories of Group A participants' assessments were indistinguishable in statistically significant terms on the remaining six quantitative measures used in this study, the Best of Group A spent substantially less time, in absolute terms, engaged in the task of identifying issues. Using the Rest of Group A as the baseline, the Best of Group A assessments allocated, on average, 54% less time to identifying issues than their similarly ranked colleagues (0.95 minutes vs 2.07 minutes)(Column 2). When the Best of Group A is used as the baseline, this statistically significant difference indicates that the rest of Group A spent 118% more time identifying issues than the former group's assessments.

This result is consistent with the expertise-related findings previously noted where time spent identifying issues was a distinguishing trait between Group A participants and all of Group C participants' assessments, and in particular their assessments involving laboured reasoning. This result also comports with the hypothesis that as a legal specialist's level of expertise increases (from Group C to Group A, and from the Rest of Group A to the Best of Group A), time spent identifying issues reduces, both in absolute terms and in some instances relative to the amount of time spent on synthesis. What was surprising is that this ability appears to continue to improve even amongst the most experienced and highest-ranked legal specialists.

(b) Group C Laboured Reasoning Against Group C Unlaboured Reasoning

Amongst Group C participants there was a similar distinction between what appeared to be better performances (legal-risk assessments involving unlaboured reasoning) and worse performances (assessments involving laboured reasoning). For the purposes of a quantitative comparison between these two groups of otherwise similarly ranked legal specialists, assessments involving unlaboured reasoning were used as the statistical baseline. The results are shown in the following Graph 5.10.

GRAPH 5.10 – Group C Laboured Reasoning Against Group C Unlaboured Reasoning Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	133	328	107	73	0	0	-23
WPM % Diff	11	0	0	0	-10	0	0

Insofar as this graph reflects differences attributable to or the result of laboured reasoning as compared to unlaboured reasoning, the differences in ROV between these two categories of legal-risk assessments at first appear to be the least interesting result. However, the fact that laboured reasoning is associated with higher ROV in terms of overall risk assessment is somewhat counterintuitive. Laboured reasoning is not slow reasoning in an overall sense (at 117 wpm vs unlaboured reasoning at 105 wpm)(Column 1). But as shown in Column 5, laboured reasoning does appear to involve lower ROV when engaging in synthesis compared to unlaboured reasoning, although both forms of reasoning have higher ROV than their overall ROV. In other words, all participants in this comparison increased their ROV when synthesising, but the Group C participants engaging in unlaboured reasoning did so more (111% vs 124% of average, overall ROV)(Column 5).

Laboured reasoning is also associated in this comparison with more than double the total time spent assessing legal risk, on average, as indicated in Column 1 (11.31 minutes vs 4.85 minutes). In terms of time spent identifying issues, the difference is even greater. Laboured reasoning is associated here with, on average, 328% more time identifying issues compared to unlaboured reasoning (4.45 minutes vs 1.04 minutes)(Column 2). Indeed, in assessments involving laboured reasoning, the

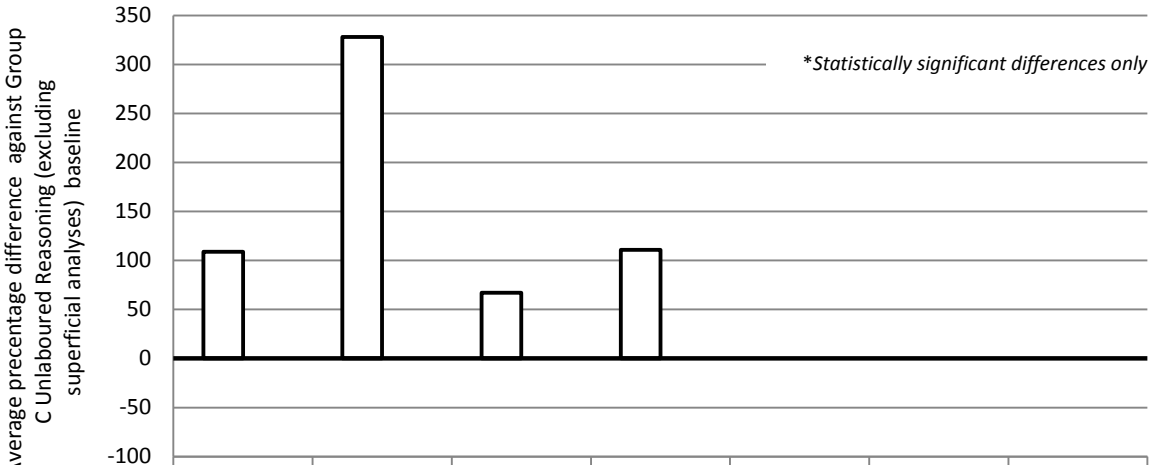
identification of issues took, on average, almost as long as the total assessment time for assessments involving unlaboured reasoning.

Twice the amount of time, in absolute terms, was also spent on synthesis by those Group C participants engaging in laboured reasoning compared to those engaging in unlaboured reasoning (2.57 minutes vs 1.24 minutes)(Column 3). This was ostensibly because their overall assessment time was on average significantly longer. As a proportion of total assessment time, laboured reasoning involved spending, on average, 72% more time identifying issues (38% vs 22%)(Column 4). Given that there was 23% less reading and clarifying time amongst those participants who engaged in laboured reasoning as a percentage of total assessment time (40% vs 52%)(Column 7), issue identification was ostensibly the principal reason why assessments with laboured reasoning took significantly longer than those with unlaboured reasoning.

(c) Group C Laboured Reasoning Against Group C Unlaboured Reasoning (Excluding superficial analyses)

Lastly, quantitative comparisons were undertaken between legal-risk assessments involving laboured reasoning and legal-risk assessments involving unlaboured reasoning but excluding assessments based on superficial analyses. As can be seen from the following graph, this resulted in a number of statistically significant differences relating to time spent during assessments overall, identifying issues and engaging in synthesis.

GRAPH 5.11 – Group C Laboured Reasoning Against Group C Unlaboured Reasoning (Excluding Superficial Analyses) Baseline*



	1. Overall	2. Identifying	3. Synthesis	4. Identifying / Overall	5. Synthesis / Overall	6. Synthesis / Identifying	7. Reading Clarifying
Time % Diff	109	328	67	111	0	0	0
WPM % Diff	0	0	0	0	0	0	0

As the exclusion of superficial analyses from these calculations affected the baseline statistics of unlaboured reasoning and not the top-line statistics of laboured reasoning, it is important to assess the resulting variations carefully. For example, excluding superficial analyses increased the overall duration of unlaboured assessments. This, in turn, reduced the relative difference between laboured and unlaboured reasoning on this measure and consequently reduced the Column 1 differential. In other words, increasing the average assessment time for unlaboured assessments (by excluding superficial analyses) made them more similar to laboured assessments and therefore reduced the recorded average difference between them, although that difference remained both substantial and statistically significant.

There were no statistically significant differences between these two categories of assessments in terms of ROV. This suggested that the differences in ROV identified when superficial analyses were included reflected the lower ROV of superficial analyses relative to other Group C assessments.

In terms of overall time taken, as superficial analyses resulted in shorter assessments, excluding them lengthened the average duration of unlaboured reasoning assessments relative to laboured ones. This resulted in a reduced, but still substantial, differential of

109% as shown in Column 1 above (11.31 minutes vs 5.42 minutes[509]). The exclusion of superficial analyses did not change the percentage differential in time spent identifying issues between laboured and unlaboured reasoning assessments. Laboured reasoning was still, on average, associated with 328% more time on this measure (4.45 minutes vs 1.04 minutes)(Column 2).

In relation to time spent on synthesis, there was a less substantial (although still statistically significant) difference between those assessments involving laboured reasoning and those involving unlaboured reasoning with superficial analyses excluded. As shown in Column 3 of Graph 5.11, the former category of assessments involved 67% more synthesis in absolute terms (laboured 2.57 minutes vs unlaboured excluding superficial analyses 1.54 minutes). This difference ostensibly reflected the fact that assessments with laboured reasoning were, on average, longer in duration. The fact that the proportion of synthesis as compared to total time was not found to be statistically significant between these two categories of assessment (Column 5) confirms this conclusion.

In addition to laboured reasoning being associated with longer legal-risk assessments compared to ones involving unlaboured reasoning excluding superficial analyses, the proportion of time spent identifying issues was also greater. After superficial analyses had been excluded, this difference was greater than when they were included. As shown in Column 4 of Graph 5.11, time spent identifying issues as a proportion of overall assessment time in laboured reasoning was more than double unlaboured reasoning (111% more) when superficial analyses were excluded (38% vs 18%). When superficial analyses were included with unlaboured assessments, laboured reasoning involved only 73% more time spent on identifying issues as a proportion of overall assessment time (38% vs 22%).[510]

Excluding superficial analyses otherwise reduced the statistically significant quantitative differences associated with laboured and unlaboured reasoning in terms of

[509] The overall duration of assessments with unlaboured reasoning inclusive of superficial analyses averaged 4.85 minutes.
[510] Column 4 of Graph 5.10.

the proportion of participants' time spent on reading and clarifying the data provided in the test documentation.[511]

(d) Summary

The comparison of the Best of Group A against the Rest of Group A baseline confirmed that there was generally no meaningful difference between these two groups on most of the quantitative measures used here. The one exception was a statistically significant difference in time spent by the Best of Group A on identifying issues. In absolute terms, even though there was no statistically significant difference between time spent overall on assessing the test cases, the Best of Group A participants spent less than half the time identifying issues compared with the baseline.[512] The magnitude of this difference was perhaps surprising. Given that this was an area where less expert participants in Group C were distinguishable from more expert participants in Group A, time spent on this task also appears to have been a reliable point of distinction between different levels of expertise at the higher end of the specialist legal expertise continuum.

With assessments involving unlaboured reasoning as the baseline (both inclusive of and excluding superficial analyses), comparisons against assessments involving laboured reasoning confirmed a common theme. Laboured reasoning by Group C participants was associated with spending, on average, more than four times the amount of time identifying issues in absolute terms (due in part to laboured reasoning leading to longer overall assessment times),[513] and more than 70% and up to 111% more time spent on identifying issues as a proportion of total assessment time compared to the baseline statistics of unlaboured Group C assessments.[514] There was not, however, a corresponding increase in synthesis time as a proportion of total time (even though laboured reasoning typically involved spending more time on synthesis in absolute terms) as compared to the unlaboured reasoning baseline, whether including or excluding superficial analyses.[515] The recorded lower overall average ROV and

[511] See the nil result in Column 7 of Graph 5.11 compared with the result in Column 7 of Graph 5.10.
[512] Column 2 of Graph 5.9.
[513] Column 2 of Graph 5.10.
[514] See Column 5 of graph 5.10 and 5.11, respectively.
[515] See the nil results in Column 5 of graphs 5.10 and 5.11.

reduced increase in ROV when engaging in synthesis compared to overall ROV only existed when superficial analyses were included in the baseline measures.[516]

D *Conclusion*

This chapter described in detail several different ways in which the verbal data elicited from participants in this study were categorised to highlight cognitive differences between higher and lower-ranked legal specialists. The ranking of participants, in terms of their levels of likely expertise in competition law was undertaken prior to the above analysis and permitted attention to be given to observed behavioural and performance differences between Group A and Group C participants, who were deemed to include either experts and masters (Group A) or apprentices and journeymen (Group C) as defined by Hoffman's Scheme of progressive, domain-specific expertise.

Qualitative differences between the performances of participants in these two groups were analysed in terms of ease of reasoning, certainty of assessment, assessment accuracy and depth of analyses. In each of these areas, test results indicated that legal specialists with different levels of expertise think differently when faced with the same risk-assessment tasks requiring the exercise of their specialist legal skills. Specifically, lower-ranked participants had an observed greater propensity for laboured reasoning, inconclusive assessments, incorrect assessments and correct assessments based on superficial analyses. Higher-ranked participants, on the other hand, engaged only in unlaboured reasoning, and provided a greater number of correct conclusive assessments, all of which relied on substantive legal or economic analyses. Only higher-ranked participants refused to provide any risk assessment in individual cases.

Quantitative differences between Group A and Group C participants were then considered with reference to verbalisation rates and time spent identifying issues and synthesising those issues in the course of assessing legal risk. An overall comparison between the performance of participants in these two groups suggested that lower-ranked participants, as a whole and on average, had lower ROV when assessing the test cases generally, and also when identifying issues and engaging in synthesis. They also spent approximately twice the amount of time identifying issues and half the time synthesising those issues compared to their allocation of time to the overall assessment

[516] Compare columns 1 and 5 of Graph 5.10 with the same columns of Graph 5.11.

task and to identifying issues. More detailed comparative results yielded further data concerning quantitative differences between these two groups in relation to several of the above areas where qualitative differences had been observed. Quantitative differences were also recorded between participants within the same expertise group. This involved comparisons between the best performances of Group A participants compared to the other performances of Group A participants, and between Group C participants who engaged in either laboured or unlaboured reasoning.

Given the objectives of the methodology underlying this study and in particular the identification of cognitive differences guided by an exploratory comparison of different assessment behaviours arising from the original research question, the results from the analyses recorded in this chapter provide a focus for the discussion in the next chapter. That discussion, which involves further consideration of the above results, marks the final stage of the study's empirical investigation.

VI ANALYSIS

The test results recorded in the previous chapter highlighted a number of differences between how legal specialists with different levels of expertise assess legal risk in an information-constrained and time-limited context. The following discussion seeks to explain these differences with the aim of providing a detailed description of cognitive traits and indicators associated with different levels of specialist legal expertise. These traits and indicators form the basis of a response to the research question, which is concerned with finding readily identifiable and measurable differences between how legal specialists with different levels of expertise think when assessing legal risk.

Part A of this chapter focuses on the time participants spent identifying issues while assessing test cases. This task was conceptualised as a cognitive activity requiring the retrieval of relevant information from a participant's Long-Term Memory ('LTM'). Based on the test results and comparisons between participants with differently levels of expertise, a statistically significant association was observed between the proportion of total assessment time spent on this task and participants' levels of expertise. This association appeared to reflect lower-level participants' less efficient and less effective retrieval of information from LTM, which in qualitative terms included spending more time recalling irrelevant and low-quality information.

Part B describes how higher-level participants, when compared to lower-level participants, spent significantly more time (as a proportion of total assessment time) drawing inferences from the information in their Working Memory ('WM'). There was no statistically significant data to link this difference to lower-level participants' greater propensity for inconclusive assessments. However, there was a meaningful association between this measure of cognitive activity and the ability of higher-level participants to base their correct assessments on the substantive analysis of relevant issues while the correct assessments of lower-level participants relied on superficial analyses.

Part C describes comparisons between the rates of verbalisation ('ROV') of participants with different levels of expertise as they engaged in the legal-risk assessment tasks used in this study. It identifies those contexts in which lower-level participants appeared to have exerted greater cognitive effort, and where higher-level experts maintained more

consistent verbalisation rates as they switched between retrieving relevant information from their LTM and drawing inferences from information in their WM.

Part D discusses the deeper cognitive processes that appear to have shaped the different ways in which study participants engaged with the test cases. This involves an analysis of the 'why' behind the 'how' of the observed identifiable and measurable differences. The theoretical base of this analysis is the work of Kahneman and others on the roles of intuition and analytical reasoning in decision-making and matters of judgment. This includes a consideration of the extent to which participants relied on one or other of these thinking types and the nature of the self-monitoring process across different levels of legal specialists.

Part E reviews other factors that could have influenced the results and inter-group analyses presented in this chapter. This review covers the potential influences of external time pressures on participants, participants' access to Internet search while assessing test cases, the novelty of the testing procedures, and the relevance of participants' prior experience with merger matters, their professional focus, whether or not they had economics qualifications, the degree to which they specialise in competition law, and their experience as staff of a competition authority.

Part F concludes the chapter with a summary of the findings and analysis presented. This summary includes a table that lists the key behavioural identifiers, performance traits and cognitive indicators associated with different levels of specialist legal expertise. It also highlights those identifiers, traits and cognitive indicators unique to a specific level of expertise. The relevance for these findings in relation to the research questions is also indicated.

The discussion in this chapter involves a return to traditional definitions for the four expertise levels of study participants. The previous chapter focused on observed qualitative and quantitative differences between Group A and Group C participants. Based on the methodology used to select and rank participants, these two groups were assumed to contain individuals at the apprentice and journeyman levels (Group C) and at the expert and master levels (Group A). Rather than continue with references to these two groups, the more descriptive terminology of Hoffman's Scheme and the traditional conceptualisation of developmental expertise are used in this chapter.

In this context, the term ‘apprentice’ refers to those Group C participants who engaged in laboured reasoning when assessing a test case. Strictly speaking, it was their assessments that were considered to be apprentice-level, as the same participant might have also engaged in unlaboured reasoning in another case, where their assessment would be considered that of a journeyman. It was the ease of reasoning evident in the assessment itself that established the appropriateness of the descriptor ‘apprentice’ when laboured reasoning was identified, and the descriptor ‘journeyman’ when it was not.

This approach assumed that each assessment was a discrete event independent of a participant’s performance in any previous or subsequent test-cases. While this introduced a complication to the analysis in terms of associating levels of expertise to individuals only indirectly, participants could still be categorised according to how often or to what degree they engaged in behaviours associated with certain levels of specialist legal expertise. For instance, participant S12 who engaged in laboured reasoning in each of the test cases he considered, could be categorised as more of an apprentice than, say, s16 who engaged in unlaboured reasoning in three of the four test-cases he considered, and may therefore be considered more a journeyman, even though he engaged in laboured reasoning in one instance.

The characteristics of laboured reasoning included the extended consideration of both relevant and irrelevant issues, difficulties with identifying key or pivotal issues, repeating earlier points of contention without resolution (in contrast to more linear problem-solving approaches) and typically longer assessment times. These risk-assessment traits were found to be strongly associated with lower-levels of specialist expertise inasmuch as 50% of all legal-risk assessments by Group C participants included laboured reasoning, while no such reasoning was observed in the assessments of Group A participants.

Assessments undertaken by Group A participants were divided into expert and master assessments on the basis that the latter category of assessments resulted in conclusive and correct opinions based on substantive legal and economic analyses. These assessments, which demonstrated most clearly the highest level of legal expertise, constituted 50% of all assessments undertaken by Group A participants in relation to

test cases A, B, C and D.[517] As noted in the preceding chapter, time spent identifying issues was the main area of statistically significant, quantitative difference between these master-level performances and those of other Group A participants, which were considered to be merely expert-level.[518]

Just as laboured assessments were the hallmark of Group C participants insofar as they were not observed amongst Group A participants, so conclusive and correct assessments based on substantive analyses were the hallmark of Group A participants. It was therefore logical to use such assessments as a means of distinguishing between more expert and less expert intra-group assessments. However, there remained an unavoidable arbitrary character to such distinctions, which was arguably more of an issue in relation to higher-ranked participants. This is because the increase in context-dependency amongst this group could be anticipated to result in even the highest level experts refusing to offer any assessment of legal risk on grounds of having insufficient information. Accordingly, inconclusive and 'no-assessment' responses could not always be assumed to be associated with less expertise compared with responses that were conclusive, correct and soundly based.

On the other hand, none of the correct assessments by Group A participants involved guessing. They were therefore compelling, as well as correct, for the right reasons. They were also demonstrations of complete responses to the representative task set for participants, and in this sense were considered optimal. Moreover, as half of Group A assessments fell into this category, comparative statistics were more likely to be meaningful than if they represented a significantly smaller or larger portion relative to the incomplete responses amongst Group A participants. For these reasons, it was considered reasonable to use such assessments as the hallmark of master-level performance, notwithstanding the likely conservative bias.

A *Identifying Issues – Retrieving Information from LTM*

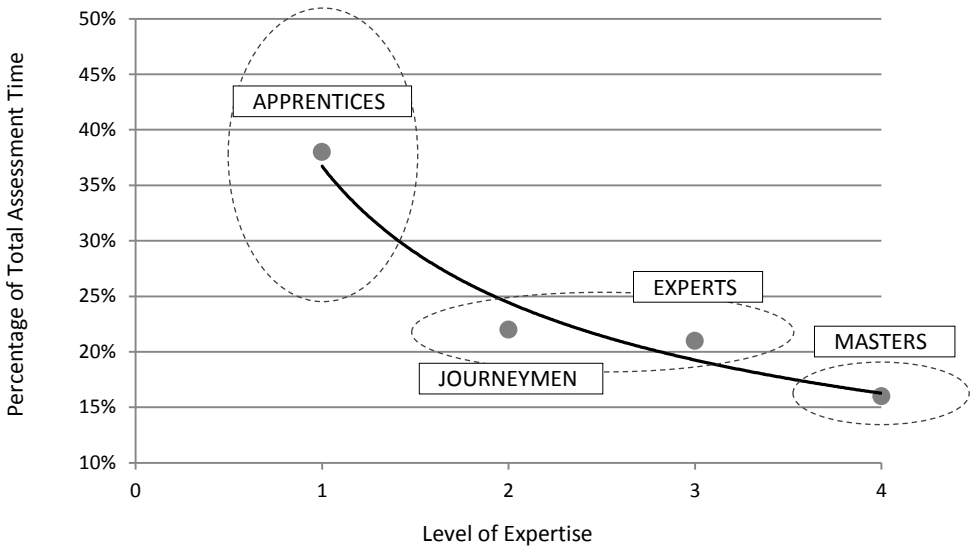
The proportion of total assessment time spent identifying issues was a clear differentiator between study participants with different levels of specialist legal

[517] Participant S01's assessments in relation to test cases E, F and G were referenced for the qualitative comparisons between Group A and Group C risk assessments, but not for the quantitative analysis described in the previous chapter. This participant's assessments are similarly used, where appropriate, for illustrative purposes in this chapter.

[518] On average, correct assessments by Group A participants involved 54% less time identifying issues as compared to the other assessments undertaken by Group A participants. See Graph 5.9.

expertise. Not only did it separate apprentices and journeymen from experts and masters, it also distinguished assessments by apprentices from those of journeymen, and those by experts from those of masters. The following chart depicts these results.

CHART 6.1 – Average Proportion of Total Time Spent Identifying Issues



These scatter points indicate statistically significant differences between different pairings of these four levels of expertise.[519] As between apprentice-level assessments (those involving laboured reasoning) and journeyman-level assessments (unlaboured reasoning), there was a statistically significant difference between the former participants' spending, on average, 38% of their total assessment time identifying issues while the latter spent only 22%.[520] As between experts and masters, the assessments of the former group involved, on average, double the amount of time identifying issues compared to the latter, but the same overall time when assessing cases.[521] While the statistical significance of the difference between these expert and master-level assessments with respect to time spent identifying issues as a proportion of total time was just outside the 95% confidence interval,[522] there was a statistically

[519] With respect to the dotted ellipses in the chart approximating the variance from the mean of individual sample points within each of the above four expertise levels, these are simply illustrative. They are not intended to indicate any specific numeric values, but rather offer a sense of the general assumptions underlying distribution around these means. The most significant aspect of these visual signifiers in this chart is the statistically distinct statuses of the apprentice and master averages, and the indistinct statuses of journeyman and expert averages as between each other.
[520] See Graph 5.10 in the previous chapter.
[521] See Graph 5.9 in the previous chapter.
[522] Student's T-Test p = 0.0657.

significant difference when words verbalised while undertaking this task were compared, which confirmed the above distinction.[523]

As between journeyman and expert-level assessments, both of which involved unlaboured reasoning, there was no evidence of a statistically significant difference between the average proportion of total time these participants spent identifying issues, whether considering time spent or words verbalised.[524]

The trend-line in Chart 6.1 approximates the above differences as a function of a participant's level of expertise. Given that these are averaged data, care must be taken in interpreting this line too literally. That said, an analysis of variance at the 95% confidence level suggests that there is explanatory power in this line-of-best-fit.[525]

1 *Discussion*

Given that laboured versus unlaboured assessments distinguished apprentices from journeymen, and correct conclusive assessments based on substantive legal or economic reasoning separated the masters from the experts, it can be contended that both these areas of difference were to a substantial degree based on, or materially related to, time spent on identifying issues.

As explained in the previous chapter, the task of identifying issues involved the cognitive processing required in Stage 2 of the effective problem-solving model described by Feltovich, Prietula and Ericsson[526] as informed by Baddely's conceptualisation of the functions of WM, short-term memory ('STM') and LTM.[527]

---

[523] All tests for statistical significance included data relating to words verbalised as well as to time spent and words-per-minute as shown in the previous graphs. This was an alternative measure used to confirm marginal statistical results. In this instance, the result for words verbalised when identifying issues as an average proportion of total words verbalised by experts and masters during their assessments were found to be from statistically distinct samples.

[524] See Graph 5.5 in the previous chapter.

[525] The analysis of variance (ANOVA) undertaken using SPSS software with respect to the power regression line in this chart resulted in a p-value of 0.026 (which was within the 95% confidence interval used in this study), and an adjusted $R^2 = 0.924$, indicating explanatory significance. One noteworthy assumption underlying this chart is that the horizontal distances between each of these levels of expertise (as indicated by the numbers 1-4 on the horizontal axis) are equidistant. Which is to say, the above chart relies on the assumption that one unit separates each of these four levels of expertise, which was an assumption that could not be confirmed within the scope of this study.

[526] Paul J Feltovich, Michael J Prietula and K Anders Ericsson, 'Studies of Expertise from Psychological Perspectives,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 41.

[527] A Baddeley, 'Short-Term and Working Memory,' in E Tulving and F Craik (eds), *The Oxford Handbook of Memory* (Oxford University Press, 2000) 77; A Baddeley, 'Working Memory: Theories, Models, and Controversies' (2012) 63 *Annual Review of Psychology* 1.

The defining characteristic of this stage is the retrieval of relevant information from LTM, from where it enters into STM/WM and is subjected to the drawing of inferences in Stage 3 of this model.

Feltovich, Prietula and Ericsson have observed that, '[i]n practice, a large part of expert problem solving is being able to access relevant knowledge, at the right time, for use in working memory.'[528] The longer time taken and, as will be noted further below, the lower ROV while attempting to retrieve such knowledge, was plausibly caused by the lower-ranked legal specialists either not having sufficient knowledge available in their LTM or not having it in as easily accessible form compared to their more expert colleagues.

The unavailability of directly relevant information in LTM seems the less plausible possibility insofar as only a small minority of participants indicated any familiarity with the industries in which the merger parties in the test cases operated. Some knew of one or other of the parties, but in other contexts or for other reasons. In this general sense, all participants were equally disadvantaged as far as direct prior knowledge on which to base an opinion on the underlying economic impact of the mergers in question. In terms of technical knowledge of the law, through the choice of representative task and the process for selecting participants – as well as from confirmatory comments made by participants themselves in the pre-test interview and during testing – there was no reason to believe that any participant had a better understanding of how the relevant laws operated either procedurally or substantively than any other participant.

It is more likely that those participants who engaged in unlaboured reasoning were able to find more effective analogical data from their LTM more quickly, which is more an issue of accessibility and information organisation. Based on cues within the test-case documentation, higher-level participants were likely better able to recognise the significance of the provided data and to match it with relevant information in their memories. They were also likely better at avoiding consideration of the kinds of irrelevant information that would increase their time on Stage 2 processes without advancing their assessment of legal risk.

[528] Feltovich, Prietula and Ericsson, above n 526, 58.

With respect to their use of analogies, whereas experts and masters used appropriate and specific analogies, apprentices tended to rely on inappropriate and vague analogies. To illustrate this point, expert and master-level participants S04, S08 and S09 consistently used directly applicable analogies and made specific references to previous cases and analogous industries. In Case A, S04 analogised with other building materials which could be substituted with one another in functional terms.[529] In Case B, S08 specifically referenced the Toll/Patrick merger and the related railway business that he correctly noted had similarities to the instant case.[530] In Case C, S04 correctly referenced the 'CHEP-type' business model as being analogous to how IBC's are pooled for hire,[531] while S08 made a similarly direct comparison with pallet hiring.[532] In Case D, S04 identified similar market-definition issues raised in several other industries, including radiology, pathology services, supermarkets, liquor stores and petrol retailing.[533] Participant S09 similarly compared the analysis of rural merchandise stores to supermarket merger analysis.[534]

Apprentice-level participants S12, S14 and S20, on the other hand, formulated poor and sometimes misleading analogies and typically used vague references to identify them. In Case C, S12 spent over 8% of his total assessment time considering the injection moulding technologies of the Pact Group, which he thought may be relevant although there was no indication in the test case documents or in the final clearance decision that this was ever considered by the ACCC.[535] In the same case, S14 described how the issues in the case were 'bringing back memories now of some work I've done in the past where – I don't know if it was IBCs, but similar,'[536] while S20 considered imports in that case to be 'quite important because I think I worked on a previous case where imports for a similar type of product were quite a significant constraint.'[537] In Case D, S12 incorrectly compared the merger parties' finance operations to those of a furniture

[529] S04A Line 140-142.
[530] S08B Line 142-149.
[531] S04C Line 41-43.
[532] S08 Line 174-178.
[533] S04D Line 72-76.
[534] S09D Line 165-168.
[535] S12C Line 56-65.
[536] S14C Line 176-178.
[537] S20C Line 65-66.

retailer[538] and S14 drew incorrect parallels between wool broking and the licensing of mortgage and insurance brokers.[539]

At the highest end of the expertise spectrum, masters needed the least amount of time to identify issues and did so with no apparent disadvantage to their overall legal-risk assessments. This may be explained by their not having to search as widely or with as much effort to locate relevant information within LTM. The information they did find was also of higher quality as evidenced by their correct assessments.

The obvious parallel is with chess grandmasters, whom De Groot found did not consider more moves than intermediate players, just much better ones.[540] In this respect, participants S06 and S01 both demonstrated near instant correct assessments based on very short response times during which they were unlikely to have undertaken a wide ranging search for relevant information, but rather were able to focus very quickly on the one or two factual issues that would prove determinative. This comports with Klein et al's Recognition-Primed Decision model in which highly expert individuals 'generate effective courses of action without having to consider more than a single option.'[541]

In Case C, S06 sought to confirm just one fact, namely, that the relevant merger transaction had already been completed. From that he was able to infer that the only possible remedy would be divestiture and that this was a highly unlikely outcome. Other participants who considered this case, specifically S12 and S19, also noted the significance of divestiture, but did not draw the same conclusion as S06 that the likelihood of this remedy was the determining factor.

In Case F, S01 only needed to read the names of the merger parties and perceive the case to be about publishing to conclude that clearance would likely be granted. His subsequent description of the reasons for this reaction revealed a general knowledge of the publishing industry and the impact of technological advances on the production and

[538] S12D Line 17-23.
[539] S14D Line 249-255.
[540] See A De Groot, *Thought and Choice in Chess* (Mouton, 1965). See also G A Klein, 'A Recognition Primed Decision (RPD) Model of Rapid Decision Making,' in G A Klein, J Orasanu, R Calderwood and C E Zsambok (eds), *Decision-Making in Action: Models and Methods* (Ablex, 1993) 138.
[541] Garry Klein, Karol G Ross, Brian M Moon, Devorah E Klein, Robert R Hoffman and Erik Hoffnagel, 'Macrognition' (2003) 3 *Human-Centered Computing* 81, 82.

distribution of published products. This was sufficient for him to form a conclusive and correct opinion on the likely outcome of the case within less than 10 seconds.

Contrast this with the laboured reasoning of S12 and S14's assessments of Case A and Case B, in which their immediate reactions[542] were that clearance would likely be problematic given information retrieved from their LTM concerning the history of competition problems in the concrete industry (Case A)[543] and the ACCC's ongoing concerns about acquisitions by businesses as big as Toll (the acquirer in Case B).[544] Participant S11 similarly stated within seconds of reading the market inquiries letter in Case B that, 'I think the ACCC is likely to have greater concerns about [this transaction] just by nature of the companies involved … and I know the ACCC has looked at these transactions fairly closely in the past involving Toll and Linfox.'[545]

The journeymen assessments of S20 in Case A and S16 in Case B, also involved immediate reactions to the nature of the industry (S20 in Case A[546]) and to the merger parties (S16 in Case B[547]). These reactions were also negative in the sense that they tended to raise the perceived level of legal risk, although neither of these participants went on to engage in laboured reasoning in either case.

The immediate, negative perceptions verbalised in these assessments reflected participants' awareness of issues relating to the industry and to the parties involved in these transactions. Moreover, the information they verbalised was readily available in their LTM from where it was accessed within the first few seconds of reading these test

[542] Immediate reactions were considered to be those reactions of participants that were relevant to the likelihood of clearance and which occurred within the first 60 seconds of their commencing reading of a test case.
[543] 'Just my initial thinking is that this industry … so, yeah there's [been some] section 46 proceedings … suggests that this industry is pretty concentrated … I know there've been other matters in this sector … so it's one that the ACCC watches closely.' S12A Line 5-13; 'So I know that the concrete industry is obviously a concentrated one in Australia and has had a, you know, history of cartel-type conduct … so the ACCC is always going to look at mergers like this pretty closely.' S14A Line 3-6.
[544] 'I haven't worked in this industry … so Toll I know is of major interest to the ACCC, after it acquired Patrick and then there was the undertakings and the split out of Asciano … the ACCC has got a big interest in everything they do.' S12B Line 3-8; 'I'm also thinking in the back of my head that the freight forwarding has been another area where cartel conduct has occurred in the past … so again that makes me think that the ACCC will look very closely at transactions that touch on this area, because they're familiar with it.' S14B Line 31-35.
[545] S11B Line 7-11.
[546] 'So, my first thought is that concrete is a product where you often get collusion … so, probably have to bear that in mind … there's often competition problems in the supply of concrete I think because geographic markets are so small … partly due to that reason.' S20A Line 3-10.
[547] 'Okay, Toll … always a Commission favourite.' S16B Line 1-3.

case documents. But their concerns were misplaced, and in over half of these instances these negative reactions were followed by laboured reasoning.

In none of the expert or master-level assessments was there evidence of similar immediate retrieval of negative (incorrect) information in relation to any test case. These participants did not verbalise such reactions, even though it can be safely assumed that they would have been at least equally as aware of the histories of these industries and the significance of the relevant merger parties in competition law sense: the concrete industry has, for instance, been of perennial interest to the ACCC (and to other competition authorities in other jurisdictions) as the above apprentices and journeymen indicated. This information was, it would seem, simply not considered relevant to bring into WM by these higher-level experts given the task at hand.

This may suggest an indiscriminate or impulsive response by mainly apprentices but also by two journeymen, who made immediate connections between what they were reading and information in their LTM. Their verbal protocols revealed that they did not filter this information for relevance, but rather acknowledged it as soon as they had accessed it. This could be confirmation of Weinstein's finding that less expert lawyers are identifiable by their imprecise and generalised use and recall of information.[548] One of the more extreme examples of such thinking was S14's 'automatic' references to possible Foreign Investment Review Board issues and merger-filing obligations in other countries as part of his initial framing of cases A and B.[549] Such matters were wholly irrelevant to the legal-risk assessment task at hand.

Apart from the immediate, positive reactions of S06 in Case C and S01 in Case F as mentioned above, the only other immediate reaction at the higher levels of expertise was by S08 in Case B where he immediately commented on the apparent narrowness of the ACCC's focus on a Bass Strait market. The nature of his reaction suggested that within less than 10 seconds of commencing to read the relevant market inquiries letter,

[548] Ian Weinstein, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1, 38.
[549] In Case A, S14 verbalised 'in Attachment A we've got Rocla – oh, okay, so it's a subsidiary of Fletcher Building, which is a New Zealand company … so I'm automatically thinking we may need to consider merger filings in New Zealand as well as Australia … and also the possibility of FIRB clearance as well.' Line 19-25; In Case B, S14 verbalised 'so, again, as soon as I hear global provider I always think about requirements for merger filings in other countries.' Line 13-14.

this master-level participant[550] had already identified the analytical significance of the ACCC's choice of a restricted geographic market and its potential to unnecessarily heighten competition law concerns.[551]

Whether these higher-level experts ignored the more general information about the parties and industries noted by the lower-level participants, or whether they simply went directly to the more determinative knowledge in their LTM without even considering this additional information, is not clear. Given the speed at which their reactions occurred, however, and the lack of any verbalised acknowledgments amongst these higher-level participants, it is feasible to propose that it was the latter process. This suggests that the mental focus of these participants precluded the triggering of excursions into related but not directly relevant contextual data held in their LTM.

An intermediate-level response may have been evident in participant S07's consideration of Case B. His assessment of this case had signs of both the apprentices' impulsive retrieval of historical information and a higher-level self-monitoring to maintain his focus on the set task. After noting the potential for the merger parties to have problems with the ACCC given their historically strained relations with the competition authority and their general prominence within their industry, S07 re-directed his focus by noting, 'but again, what you need to do is to go through and just see what this letter is indicating in terms of what it is that is being sold.'[552] He nevertheless went on to conclude incorrectly that the proposed transaction would likely be opposed.

With respect to those higher-level participants who refused to provide an assessment because of a lack of available information, a likely explanation is that these experts were more contextually dependent than less expert participants. More specifically, participant S06 who had performed at the master-level in Case C verbalised in Case D that the assessment task was impossible for him to complete without having more information. Participant S09, who was the highest-ranked participant in the study, similarly stated that he was unable to provide an assessment without satisfying his need

[550] All of participant S08's legal-risk assessments were judged as being master-level insofar as they were correct when compared against the actual outcomes of the cases and they relied on substantive legal and economic analyses.
[551] 'Oh, it's a very narrow market.' S08B Line 3. This participant went on to elaborate later in his assessment, 'I must say that freight forwarding narrowly defined in just Tasmania and the Mainland seems a bit, a bit silly.' S08B Line 93-94.
[552] S08B Line 18-19.

for further information regarding production processes (Case A) and internal sales records (Case D), amongst other requirements. Instead of providing inconclusive assessments, these two participants consistently refused to provide any opinion.

Participant S07 also refused to provide an opinion with respect to Case A, but this appears to have been a reaction to the novelty of the testing process itself. He subsequently went on to provide a conclusive – but incorrect – assessment in respect of Case B, a correct assessment based on superficial analysis in Case C and an inconclusive but unlaboured assessment in Case D. Neither S06 nor S09 could be persuaded to state even a qualified view, except for S06's initial and correct response in Case C where he immediately identified determinative information from the context of the case.

2 *Response to the Research Question*

Laboured or apprentice-level reasoning involved more time identifying issues both in absolute terms and as a proportion of total assessment time. This was ostensibly because these participants did not screen out irrelevant information as much or as effectively as higher-level participants. The highest-level participants may not have even needed to undertake such screening because their focus precluded the consideration of such information in the first place. This further reduced their time spent identifying issues. In general terms, apprentice-level assessments lacked screening, while journeyman and expert-level assessments engaged in screening – but there was no evidence that master-level assessments involved even the triggering of irrelevant or non-determinative information in LTM which would require screening.

Correct conclusive assessments observed during this study may be explained by master-level participants retrieving higher quality information from their LTM. Expert-level assessments involved more than twice as much time identifying issues compared to master-level assessments, which may have been because they spent more time considering the relevance of information and searching for determinative information. Master-level assessments involved less time identifying issues in absolute terms and as a proportion of total assessment time – yet their assessments were of higher quality (insofar as they were conclusive and correct) ostensibly because they relied on the retrieval of less information which was determinative. This behaviour may be explained by their better and more detailed risk-assessment schema which made clear

what information was required and provided cues as to where to find this information in LTM through more effective analogising.

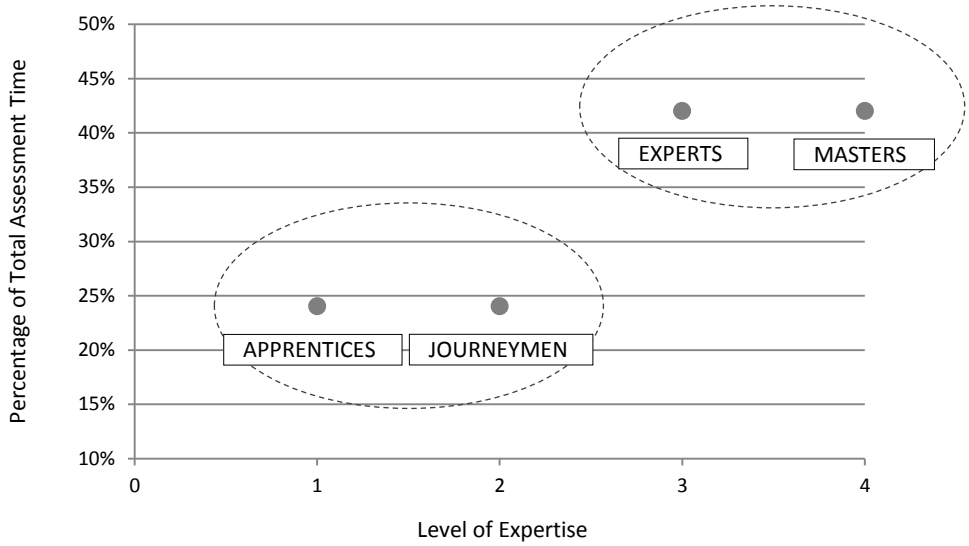
Expert-level participants who concluded that they did not have access to determinative information refused to provide an assessment. This suggested that their risk-assessment objective was a conclusive view, whereas lower-level participants were content – or at least could be persuaded – to provide inconclusive assessments. Moreover, these experts were highly discriminating and prescriptive as to the kinds of further data they required, which supports the view that they had a clearer focus on what was important (a benefit of more detailed and effective problem-solving schema), and that the quality or determinative-potential of that information was crucial to their deliberations. In other words, they would rather not waste their time considering a case without the information they believed was necessary to provide a conclusive view, which was the only objective they considered.

B *Synthesis – Drawing Inferences*

Inferring the level of legal risk in a test case or the types of further information required to assess such risk, were considered acts of synthesis within Stage 3 of the model of effective problem-solving. In this final stage of the problem-solving process, information drawn from LTM is combined with data gathered from the environment and stored in STM to enable inferences to be drawn from WM to achieve the ultimate objective of the problem-solving task, which in this study was to assess the likely outcome of a merger clearance application.

The following chart shows how the average proportion of total assessment time spent engaging in synthesis was different depending on participants' levels of expertise.

CHART 6.2 – Average Proportion of Total Time Spent Engaging in Synthesis



Adopting the same approach as Chart 6.1, this chart describes the statistically significant differences between apprentices and journeymen on the one hand, and experts and masters on the other, with respect to how much time they spent, on average, on synthesis as a proportion of total assessment time. The indicated differences are based on data presented in the previous chapter and recorded in graphs 5.3, 5.4, 5.5, 5.9 and 5.10. The result is a chart in two parts depicting the separation of higher level experts from lower level ones, but not between participants within these two groups.

1 *Discussion*

Legal-risk assessments by apprentices and journeymen both had statistically indistinguishable and a significantly lower proportion of synthesis time to total time compared to experts and masters (an average 24% vs 42%). This suggests that it was not a measure with as strong an association with laboured reasoning as time spent identifying issues. As previously shown in Graph 5.10, apprentices, who were defined as Group C participants who engaged in laboured reasoning, spent more than three times as long as journeymen identifying issues in absolute terms, and 73% more as a proportion of total assessment time. This indicated that laboured reasoning involved more time identifying issues.

With respect to synthesis, however, there was no statistical difference between these two Group C sub-groups as shown in the above Chart 6.2. Together they spent 45%

less time than experts and masters on this Stage 3 task. Less time spent synthesising was something that these lower-level participants had in common with each other, but not with the other two groups.

Another common trait or attribute was the higher rate of inconclusive assessments which was observed amongst apprentices and journeymen, but not amongst experts and masters. Overall, the former two groups had a combined 70% incidence of inconclusive assessments compared with the latter two groups' combined 25% incidence of inconclusive assessments. Amongst masters, there were no inconclusive assessments as their assessments were by definition both conclusive and correct.

Other shared traits amongst lower-level participants were (i) correct assessments based on superficial analyses, which occurred in four instances amongst apprentices and journeymen (17% of all Group C assessments), and (ii) incorrect assessments, which also occurred four times amongst this group (17% of all their assessments) and only once amongst the experts (a 6% incidence).

Less time synthesising as a proportion of total time was therefore viewed as a possible causal or symptomatic factor of inconclusiveness, superficiality in forming a conclusive assessment, and inaccuracy in assessing legal-risk.

Inconclusive assessments by apprentices and journeymen involved, on average, less than half the time spent on synthesis as a proportion of total assessment time compared to all the assessments by experts and masters (20% vs 42%).[553] However, when inconclusive assessments and conclusive assessments by apprentices and journeymen were compared to each other, there was no statistically significant difference in this area.[554] Neither was there any difference between their assessments in terms of time spent identifying issues as a proportion of total assessment time.

Statistically, therefore, it could not be concluded that less time spent on synthesis as a proportion of total assessment time was any more strongly associated with inconclusive assessments than more time spent on identifying issues. In other words, inconclusive assessments could not be distinguished from the conclusive assessments of similarly

[553] These results were not presented in the previous chapter, but were recorded during the study. The relevant Student's T-Test p-value = 0.0019 indicated that the difference between these samples on the measure of proportion of total assessment time spent on synthesis was statistically significant.
[554] This intra-group assessment result was also not presented in the previous chapter, but was recorded to during the study.

ranked participants in terms of time spent on synthesis. Accordingly, such assessments could not be said to be uniquely associated with (or caused by) less time spent on synthesis.

Superficial assessments, however, were found to be strongly associated with less time spent on synthesis, in both absolute and proportional terms. Compared to experts and masters, journeymen who undertook superficial assessments engaged in 77% less time synthesising in overall terms, and 60% less time as a proportion of total assessment time. Further, their time spent on synthesis as a ratio of time spent identifying issues was 83% less. At the same time, there was no statistically significant difference between these superficial assessments and those of experts and masters in relation to time spent identifying issues, either in absolute or proportional terms.[555]

Compared to assessments with laboured reasoning, superficial assessments (which were a sub-category of unlaboured reasoning) involved, on average, less time on synthesis overall (0.63 minutes vs 2.57 minutes) because such assessments were, on average, significantly shorter (3.72 minutes vs 11.31 minutes). In terms of time spent engaged in synthesis as a proportion of total assessment time, there was no statistically significant difference between assessments with laboured reasoning and superficial assessments.[556]

Compared to the other assessments involving unlaboured reasoning, superficial assessments also involved less time on synthesis in absolute terms because they were shorter assessments. However, there was no statistically significant difference between these two types of assessment when comparing average time spent on synthesis as a proportion of total assessment time.[557]

Given that superficial assessments by the journeymen in this study involved, on average, significantly less time spent on synthesis as a proportion of total assessment time as compared to experts and masters, but not in comparison to other assessments by apprentices and journeymen involving unlaboured and laboured reasoning, it is plausible to conclude that superficiality of analysis is associated with proportionally less time spent synthesising issues.

[555] These results were presented in Graph 5.7 in the previous chapter.
[556] These results were not presented in the previous chapter, but were recorded during the study.
[557] These results were not presented in the previous chapter, but were recorded during the study.

This is supported by Graph 5.7 which showed that the only statistically significant areas of difference between superficial assessments and the substantive assessments of higher-level experts – apart from less total assessment time – was less time spent on synthesis. At the same time, there was no less time spent on identifying issues.

Whereas laboured reasoning was considered in the previous section to be associated with more time spent identifying issues (a Stage 2 problem-solving issue), superficiality therefore appears to be associated with less time spent synthesising issues (a Stage 3 issue). In other words, superficiality is not a matter of identifying fewer issues or doing so less thoroughly, but in truncating the synthesis or inference stage of the legal-risk assessment process. Moreover, this appeared to be a reliable point of difference between apprentices and journeymen as a group compared to experts and masters as a group, as evidenced by the consistently lower levels of synthesis observed across all of the former groups' assessments.[558]

The incidence of assessments based on superficial analyses was greatest in relation to Case C, which atypically concerned a completed merger transaction. Participants S20, S11 and S16 all based their assessment of legal risk in this case on the circumstantial fact that the parties had not sought prior clearance approval, and therefore that clearance would not be a problem. No expert or master-level participant relied on this fact. Rather they engaged in substantive analysis to reach a correct conclusion.

Why lower-level experts had a greater propensity for superficial analyses is not clear. One possibility is that they consciously or subconsciously sought out short-cuts to minimise the need for synthesis, or they may have simply adopted a more opportunistic mindset. Or they might have considered more time spent identifying issues to be a better problem-solving strategy. Experts and masters, on the other hand, were more focused on synthesis at the expense of identifying issues (and they did not react to superficial shortcuts). Their assessments were therefore more likely to be built on stand-alone, reasoned analysis and to be more consistently effective.

[558] See: Graph 5.3 (Group C participants overall and on overage spent 43% less time synthesising as a proportion of total time compared to Group A participants); Graph 5.4 (Group C laboured reasoning involved 45% less time on synthesis as a proportion of total time); Graph 5.5 (Group C unlaboured reasoning involved 38% less time on synthesis as a proportion of total time); and Graph 5.7 (Group C superficial assessments involved 60% less time on synthesis as a proportion of total time).

These results may reflect the findings of previous studies that experts engage in a deeper conceptual analysis of problems than novices – and this may blind them to superficial short-cuts. Such a mindset may be necessary to achieve the more efficient, effective and substantive strategies that the highest-level experts use to solve problems. It may also indicate a greater confidence and willingness to engage in synthesis, even when information is constrained and time is limited.

An example of the more efficient use of information was participant S06's ability to decide Case C correctly with just one piece of contextual information, namely, that the only remedy available to the ACCC in this case was court-ordered divestiture, which he considered was a highly unlikely outcome. Both S12 and S19 similarly noted that divestiture was the only remedy available, while S12 further recalled that there had been no previous successful divestiture actions by the ACCC against completed mergers. Participant S19 interpreted the ACCC's interest in the case as a sign of unusual vigilance that he believed 'does sharpen the mind.'[559]

In this instance, S06 was able to immediately engage in synthesis and provide a correct and substantively reasoned legal-risk assessment within less than one minute. After noting the same contextual information, S12 by contrast embarked on a 14-minute, inconclusive assessment in which he spent more than double the amount of time identifying issues than engaging in synthesis. Participant S19 similarly provided an inconclusive assessment after more than 14 minutes of further analysis involving almost 50% more time identifying issues than engaging in synthesis.

An example of deeper conceptualising during synthesis in Stage 3 of the problem-solving process was S09's identification of required information and where to look for it. In Case A, this Group A participant inferred that it was necessary to understand more about the manufacturing process for the reinforced concrete products produced by the merger parties, and that this information would best be sought directly from personnel working in the manufacturing plants. In Case D, he inferred that the scope of the geographic market around each rural merchandise store in that case could be most effectively established from the sales records of the parties. Other less expert participants had either not understood the key issues in these cases or were only

[559] S19C Line 25.

capable making a general complaint on the need for more – but largely unspecified – industry information.

In summary, those participants who spent more time engaging in synthesis as a proportion of total assessment time appeared to have had a clearer idea of what information was most important and how to best to get it. When they found determinative data, they correctly identified it as such and then stopped seeking to identify further issues. Their focus was, in this sense, on synthesis and the inferring of propositions likely to lead to a conclusive – and correct – legal-risk assessment.

The accuracy of participants' assessments was the third area where the reduced emphasis on synthesis amongst apprentices and journeymen may have some explanatory significance. This is because the incidence of incorrect conclusive assessments was much higher amongst these participants than amongst experts and masters, who also engaged in more synthesis. An obvious issue was therefore whether there might be a connection or an association between less synthesis and incorrect assessments.

Comparisons between, on the one hand, assessments by apprentices and journeymen that resulted in incorrect opinions on the level of legal risk in a particular test case and, on the other hand, the assessments of experts and masters (only one of which was incorrect), suggested that time spent on synthesis was not a distinguishing factor. Rather, it was the relative rates of verbalisation when synthesising compared to overall ROV and the ROV when synthesising compared to the ROV when identifying issues, where differences were observed.

Graph 5.8 in the previous chapter showed that there were no statistically significant differences between the above two groups in terms of synthesis time, either in absolute or proportional terms. But it did show that, on average, participants producing incorrect assessments verbalised 12% slower overall, 25% slower when identifying issues and 11% slower when engaging in synthesis. At the same time, they sped up more when synthesising compared to their overall verbalisation rate, and their ROV when synthesising was 22% higher as a ratio of their ROV when identifying issues as compared to experts and masters.

As these were the only noted differences, and because the next section focuses exclusively on differences in verbalisation rates, these issues are discussed further below.

2 *Response to the Research Question*

Inconclusive assessments could not be directly associated with less time spent synthesising as there was no statistically significant difference on this measure between the inconclusive and conclusive assessments of apprentices and journeymen. This is notwithstanding that overall, apprentices and journeymen engaged in synthesis significantly less as a proportion of total assessment time when compared to experts and masters.

Correct assessments by journeymen based on superficial analyses appeared to be associated with less time spent synthesising, both as between such assessments compared to other journeymen and apprentices, and compared to experts and masters. In fact, superficial assessments were the most different from experts and masters-level assessments in respect of this measure. At the same time, these journeymen were indistinguishable from higher-level experts and masters as a whole in terms of time spent identifying issues. Reduced synthesis therefore appeared to be causally related to or a symptom of superficial analysis.

One possible explanation for this is the more serious approach experts and master-level participants had to conceptualising the issues and to crafting their own stand-alone analyses. This may have prevented them from taking side-trips into possible superficial assessments, even when they explicitly noted that the circumstances for such assessments exist.[560] If so, this may be seen as another aspect of the more detailed schema that also prevents them from engaging in the wide-ranging identification of issues associated with laboured reasoning as discussed in the previous section. Meanwhile, less expert legal specialists remain alert to any opportunities to avoid engaging in synthesis insofar as this task is more difficult for them than for their more expert colleagues who are more capable of forming their own reasoned opinions.

[560] For instance, S04 recognised in his assessment of Case C that the merger parties had completed their merger without seeking ACCC clearance, but he then went on to complete a thorough analysis of product substitution and barriers to entry. Similarly, S05 explicitly reiterated that this case involved a completed transaction, but made no further comment on this fact, while S08 simply noted 'completed transaction … okay, so it's already happened,' and then embarked on his longest analysis of the four cases he correctly assessed (S08C Line 9-11).

Experts and masters may also have the advantage of being able to synthesise more efficiently and effectively with limited information. For example, S06 was able to base his reasoning as to why the transaction in Case C would not be opposed merely on the fact that divestiture was the only remedy available to the ACCC. Other less expert participants noted the same point, but only S06 could discern the deeper significance of this fact in the context of the legal-risk assessment task. He did not embark on the further identification of issues, because this was unnecessary. It was also unnecessary for him to reflect on the fact that the parties had not sought prior clearance, which would ostensibly have taken more time to consider and result in a less robust assessment in terms of substantive analysis.

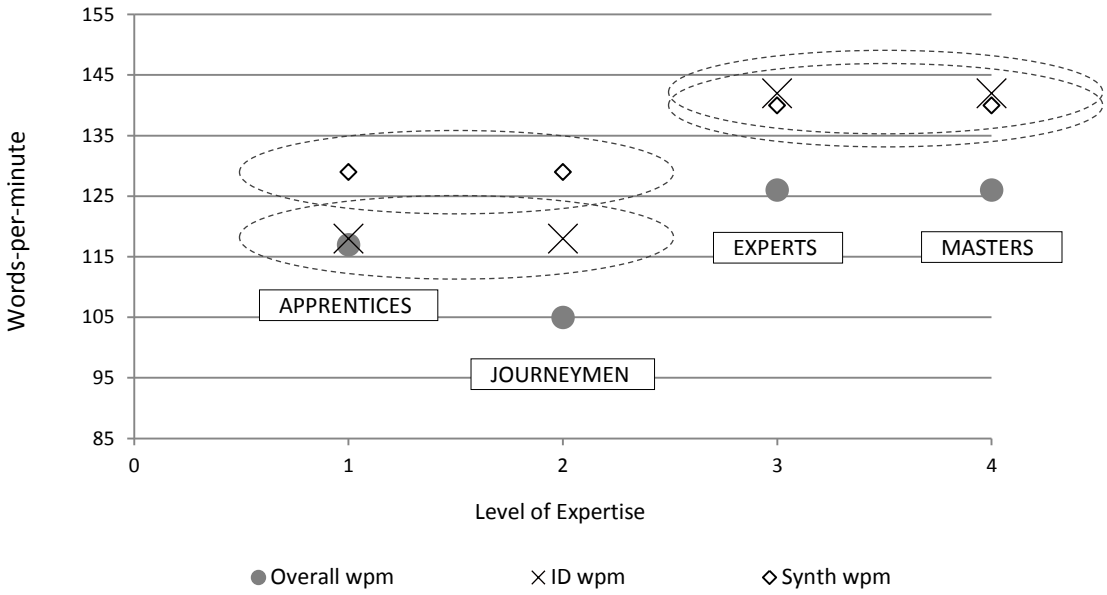
It is also possible that by being better at identifying issues, experts and master-level participants lay the groundwork for easier and more efficient synthesis and drawing of inferences. Because they know or can infer what information is most relevant and useful, they reduce their effort identifying issues. This also means that they have higher-quality information to work with when engaging in synthesis. Both these traits may be traced back to a more effective mental schema that remains operational even (or, perhaps, especially) when time is constrained and available information is limited.

C *Verbalisation Rates – Cognitive Load*

Verbalisation rates, which are measured as words-per-minute ('wpm'), provide information about the cognitive load or thinking difficulty that a participant is experiencing while solving a problem. In simple terms, the higher a participant's ROV the less cognitive capacity is being used as evidenced by their still being able to freely verbalise their thoughts. Low ROV is conversely associated with the use of greater mental effort to the point that verbalisation becomes more difficult and less fulsome. Previous studies and theoretical support for these relationships and associations between ROV and cognitive activity were discussed above in Part C of Chapter 5.

The following chart shows the average ROV for all four levels of study participants in relation to their overall assessment of test cases, their identification of issues (Stage 2 of effective problem solving) and their engaging in synthesis (Stage 3).

CHART 6.3 – Average Verbalisation Rates



1 *Discussion*

As with previous charts, the markers on this chart indicate statistically significant differences between participants according to their expertise category. For the most part, the ROV of apprentices and journeymen were assessed as being equal. The exception was that in overall terms, assessments by journeymen had a lower average ROV than assessments by apprentices (105 wpm vs 117 wpm).[561] Expert and master assessments were found on average to have equal ROV – and substantially higher ROV than apprentices and journeymen – on all measures. This suggested that participants in the two higher-level expert categories experienced less cognitive difficulty undertaking the given legal-risk assessment tasks compared to those in the two lower-level expert categories. Moreover, this was the case in terms of assessments overall as well as in relation to the identification and synthesis of issues.

Chart 6.3 also shows that the ROV of experts and masters was very similar whether they were engaged in Stage 2 or Stage 3 analysis, that is, whether they were retrieving information from their LTM or drawing inferences in WM. The ROV of apprentices and journeymen, however, was different depending on their stage of assessment. These participants' ROV when drawing inferences (engaging in synthesis) was higher than their ROV when retrieving information (identifying issues). This suggested that the

[561] See Graph 5.10 in the previous chapter.

latter task was more taxing in terms of cognitive effort than the former task, whereas between experts and masters there was not as noticeable a difference.

After reviewing the results of various inter-group and intra-group comparisons with respect to ROV as discussed in the previous chapter, the general finding was that Stage 2 tasks were more difficult – in a cognitive load sense – for lower-ranked participants. In most instances, the lower ROV's for apprentices and journeymen when identifying issues was more than could be explained by their lower overall ROV. Their lower ROV when synthesising, as compared to the ROV of experts and masters when synthesising, was more easily explained as a function of their simply having a lower ROV across their assessments generally. In those instances when their ROV for synthesis was higher relative to their overall ROV, it did not follow that their ROV for synthesis was higher in an absolute sense, but only relation into their comparatively lower overall ROV.

One group of journeymen, however, increased their ROV when synthesising compared to their overall ROV, which was statistically the same as for the higher-ranked participants. These were the participants who engaged in superficial analyses. Not only did they spend less time engaging in this Stage 3 task compared to experts and masters as noted in the previous part, they also sped up more compared to their overall average.[562] This further suggested that this difference was integral to or at least a feature of their superficial approach. In effect, their synthesis of issues required less cognitive effort precisely because it was so superficial, that is, shorter and less taxing.

Compared to the baseline assessments of experts and masters with respect to their ROV when identifying issues: assessments with laboured reasoning (apprentice-level assessments) involved 13% lower ROV; assessments with unlaboured reasoning (journeyman-level assessments) involved 20% lower ROV on this Stage 2 task; superficial assessments (a subset of journeyman-level assessments) had ROV that were statistically no different; and incorrect assessments had 25% lower ROV.

Compared to the baseline of assessments with unlaboured (journeyman) reasoning, however, none of the above intra-group differences was statistically significant. In other words, while there appeared to be differences between these Group C

[562] See Graph 5.7 in the previous chapter. Their rate of increase from overall ROV to ROV while synthesising was 17% higher than for experts and masters.

subgroupings when compared against the Group A baseline, none of these differences were statistically significant inter se. The only exception was the previously noted (and depicted in the above chart) lower overall ROV for assessments with unlaboured reasoning (journeyman assessments) compared to assessments with laboured reasoning (apprentice assessments).

The inconclusiveness of these comparisons may have arisen from small sample sizes, particularly given that intra-group comparisons involved a total population of no more than a third of all samples recorded in the study. Accordingly, larger sample sizes may have revealed more and greater differences at the required level of statistical confidence.

Another factor was the process of counting words for each 10-second interval categorised as either identifying issues or synthesising issues. Reading and clarifying was not subjected to word-counts directly because, as previously noted, such differences would be too easily influenced by differences in individual reading styles – some of which involved high rates of verbalization while others involved silent reading.

Instead of counting letters and then calculating 5-letter words as Deffner had done within his 4-second intervals,[563] whole words were counted for this study, which inevitably meant that the lengths of words could materially affect the words-per-minute calculation performed for each 10-second interval. This lack of precision, while unavoidable within the generalised nature of the present approach, arguably reduced the reliability of statistical comparisons based on these data. This was less a problem for overall ROV, but it was likely to have had an impact on the smaller-sample comparisons at the task-by-task level.

That said, there was sufficient statistical robustness in two key areas to permit the formulation of some useful propositions relating to this thesis' research question and the reasons behind the qualitative differences noted in the previous chapter.

2 *Response to the Research Question*

Overall, apprentices and journeymen appeared to have had to work harder than experts and masters to perform the legal-risk assessment tasks presented in the test cases. More

[563] G Deffner, *Think Aloud: An Investigation of the Validity of a Data Collection Procedure* (PhD Thesis, University of Hamburg, 1983).

particularly, they expended greater cognitive effort while retrieving information from their LTM and while engaging in synthesis. This may explain the laboured nature of some assessments, the inconclusive nature of others, and why more of their assessments were incorrect. It may also provide insights into why some of these participants could have been motivated to seek out superficial answers, while experts and masters, who had the capacity to engage in substantive analyses with little apparent increase in cognitive load, had no such need.

Those journeymen who based their correct assessments on superficial analyses were shown to increase the ROV of their synthesis at a greater rate than more expert participants. This ostensibly reflected the former group's success in making synthesis easier for themselves by identifying superficial and circumstantial information that obviated the need for more substantive analyses. This supports the earlier suggestion that journeyman-level participants actively search for such short-cuts to their analysis, while experts and masters either ignore or are blind to them.

Lastly, with respect to journeymen's apparent lower overall ROV compared to apprentices (but the same ROV as apprentices when identifying and synthesising issues), this is logically explained by their verbalising less while reading and clarifying. As reflected in Graph 5.10 in the previous chapter, unlaboured reasoning involved spending, on average, 52% of total assessment time to reading and clarifying, whereas laboured reasoning involved only 40% of total assessment time on this task.[564] It therefore follows that a lower verbalisation rate while reading would lower journeymen's overall ROV while their ROV when identifying and synthesising issues remained the same as for apprentices.

This may indicate that journeymen read the test-case documentation more intensively, perhaps looking for data on which to base a conclusive view, even (or especially) if this involved superficial analyses. Laboured reasoning, on the other hand, involved apprentices getting lost in their own thoughts, and more specifically expending greater effort on trying to retrieve relevant information from their LTM.

[564] See Column 7 in Graph 5.10.

D *Intuition and Analytical Reasoning*

A different but complementary approach to analysing these data involves applying a conceptual framework developed by Kahneman and others[565] that distinguishes between two types of cognitive processes first labelled by Stanovich and West as System 1 and System 2 thinking.[566] Kahneman has popularised this framework in his book *Thinking, Fast and Slow*.[567] In simple terms, the distinction between these different thinking types is that System 1 (intuition or thinking fast) relies on intuitive insights and System 2 (analytical reasoning or thinking slow) involves judgments based on analytical reasoning.

Kahneman describes System 1 cognition as 'typically fast, automatic, effortless, associative, implicit (not available to introspection), and often emotionally charged.'[568] System 2 cognition, on the other hand, is distinguished by reasoning that is, in comparison to System 1 thinking, 'slower, serial, effortful, more likely to be to be consciously monitored and deliberately controlled.'[569] In terms of the relationship between these cognitive processing types, System 2 thinking is present in all judgments, whether or not such judgments begin with intuitive impressions generated by System 1 thinking.[570]

System 2 cognition is associated with self-monitoring and with the qualities of overt behaviour.[571] When these functions of System 2 thinking leave an initial System 1 impression unchanged, the final decision or solution is said to be intuitive.[572] Kahneman anthropomorphises the relationship between these two types of thinking

---

[565] See A Tversky and D Kahneman, 'Extensional vs Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment' (1983) 90 *Psychological Review* 293; S Epstein, 'Integration of the Cognitive and Psychodynamic Unconcious' (1994) 49 *American Psychologist* 709; K R Hammond, *Judgment Under Stress* (Oxford University Press, 2000); L L Jacoby, 'A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory (1991) 30 *Journal of Memory and Language* 513; S Chaiken, S and Y Trope (eds), *Dual-Process Theories in Social Psychology* (Guilford Press, 1999); R M Hogarth, *Educating Intuition* (University of Chicago Press, 2001); D G Myers, *Intuition: Its Powers and Perils* (Yale University Press, 2002).

[566] Keith E Stanovich and Richard F West, 'Individual Differences in Reasoning: Implications for the Rationality Debate?' (2000) 23 *Behavioral and Brain Sciences* 645.

[567] Daniel Kahneman, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011).

[568] Daniel Kahneman, 'A Perspective on Judgment and Choice: Mapping Bounded Rationality' (2003) 58(9) *American Psychologist* 697, 698.

[569] Ibid.

[570] Ibid 699.

[571] D T Gilbert, 'Inferential Correction' in T Gilovich, D Griffin and D Kahneman (eds) *Heuristics and Biases* (Cambridge University Press, 2002) 167; Stanovich and West, above n 566.

[572] Kahneman, above n 568, 699.

with the observation that 'System 2 monitors the activities of System 1.'[573] But he also states that such monitoring can be lax even when effortful[574] and does not always result in a better result than intuition alone may produce.[575] On this latter point, Dijkstra, van der Pligt and van Kleef have more recently provided evidence that in some contexts – particularly when intermediate-level subjects are tested – an individual's performance may be improved by reducing the influence of System 2 reasoning on otherwise correct intuitive responses.[576]

Based on the data generated in this study and the preceding analysis of the reasoning behaviours and problem-solving effectiveness of study participants, there are two areas where further inferences may be drawn in relation to expertise-related cognitive differences amongst legal specialists. These concern participants' bias towards or greater reliance on System 1 or System 2 thinking, and the effectiveness of their System 2 self-monitoring and control functions.

1 *Reliance on System 1 or System 2 Thinking*

The design of the representative task used in this study effectively precluded successful completion using System 2 reasoning alone. This is because insufficient information was provided to participants for them to undertake a fully-reasoned assessment of legal risk. The test-case documents were essentially just lists of questions directed at individuals familiar with the merger parties and relevant industries. The only inferences that could be drawn from these documents were those based on facts that were provided in the brief background summaries and those that were incidental to – or which could be deduced from – the questions themselves.

For these reasons the initial expectation was that the most competent participants in this study would provide inconclusive but unlaboured assessments. Yet some participants were able to provide correct assessments based on substantive analyses, while others

[573] Ibid.
[574] D Kahneman and S Frederick, 'Representativeness Revisited: Attribute Substitution in Intuitive Judgment' in T Gilovich, D Griffin and D Kahneman (eds), *Heuristics and Biases* (Cambridge University Press, 2002) 49.
[575] Kahneman, above n 568, 699.
[576] Koen A Djkstra, Joop van der Pligt and Gerben A van Kleef, 'Deliberation Versus Intuition: Decomposing the Role of Expertise in Judgment and Decision Making' (2013) 26 *Journal of Behavioral Decision Making* 285, 291. Klein has also noted that experts often perform better when they trust their instincts instead of engaging in reasoned analysis. See G Klein, *Intuition at Work: Why Developing Your Gut Instincts Will Make You Better at What You Do* (Doubleday, 2003) Chapter 3.

provided incorrect assessments, and still others provided inconclusive and laboured assessments, while several provided correct assessments based on superficial analyses. These outcomes, and their distribution across the expertise spectrum, may be explained in terms of individual participants' reliance on either System 1 or System 2 cognition as reflected in the data analysed in this chapter.

The key to this analysis is the incidence and reliance on cues in the test documents as to what the likely outcome of a case would be, and how such cues were subsequently processed by each participant. In this regard, Kahneman describes five different ways in which a judgment (or in this case a legal-risk assessment) may be made:

1. An intuitive judgment or intention is initiated [System 1 thinking], and
   (a) Endorsed by System 2;
   (b) Adjusted (insufficiently) for other features that are recognised as relevant;
   (c) Corrected (sometimes overcorrected) for an explicitly recognized bias; or
   (d) Identified as violating a subjectively valid rule and blocked from overt expression.
2. No intuitive response comes to mind, and the judgment is computed by System 2.[577]

Intuitive judgments were observed in the transcripts of participants at two points within the legal-risk assessment process. First were those that occurred within the first minute of reading the test case documentation. Amongst the master-level participants these occurrences often related to the ultimate outcome of a case, rather than just individual issues. They were also rarely adjusted or changed when subjected to System 2 monitoring. These judgments, which were proven correct in master-level assessments,[578] were therefore of the above 1(a) or (b) type. In terms of the previously discussed data, this method of reasoning was associated with minimal time identifying issues in terms of Stage 2 problem-solving.

The initial intuitive responses of apprentices and journeymen, however, were either incorrect[579] and/or were changed during the course of System 2 reasoning.[580] This

[577] Kahneman, above n 568, 717.
[578] Assessments in which a correct initial framing and assessment of a case was evident were (Master level) S01F and S06C.
[579] Assessments in which an incorrect initial framing of a case was evident were (Apprentice level) S12A, S14A, S14B and S11B (Journeyman level) S12B, S20A and S16B.

behaviour was consistent with options 1(c) and (d) outcomes, which reflect a greater reliance on System 2 reasoning over System 1 intuition inasmuch as System 2 reasoning prevails as the principal mediator of the verbalised legal-risk assessment. Typically, however, these intuitive judgments were more in the form of observations that were only tentatively held and 42% of the time these participants corrected their initial assumptions.

A second type of intuition occurred throughout the legal-risk assessment process, often as component inputs to the reasoning process. These were the minor insights and informed-guesses. Insofar as they involved the identification of relevant and irrelevant issues, they were part of and ostensibly contributed to the extended time that apprentices spent on Stage 2 problem-solving. But this type of intuition also occurred during Stage 3 of the problem-solving process in which participants were likely to have been switching between System 1 and System 2 thinking as they raised and then assessed different intuitive insights.

Where only few or no intuitive responses were forthcoming, participants were forced to rely on System 2 processing (Option 2 above). Inasmuch as lower-ranked participants were less likely to generate intuitive insights, they had to work harder at System 2 reasoning (which is inherently more effortful) and therefore needed to rely more on insufficient data from the provided test-case documents. But these participants also had a tendency to persevere in this task and to consider contradictory and inconclusive information. This resulted in laboured reasoning and inconclusive results, which were hallmarks of apprentice-level assessments. Journeymen and experts, however, limited their recall of information from LTM ostensibly to avoid laboured reasoning once they had determined that simply identifying more issues would not yield a conclusive result.

Spending an extended period trying to reason through the legal-risk assessments in this study arguably reflected an overreliance on intelligence over experience. Stanovich and West have shown that the ability to engage in System 2 processing is correlated with intelligence,[581] but this is likely to be attenuated by having more experience. It is therefore likely that over-confidence was at least partly the cause of lower-ranked

[580] Assessments in which an initial framing of a case was materially changed during the course of the participant's subsequent consideration of the case were (Apprentice level) S12A and S14B (Journeyman level) S20A.
[581] Stanovich and West, above n 566.

participants spending more time trying to plough through the assessment process and in particular the Stage 2 retrieval of information from LTM. This lack of inspired searching was ostensibly also reflected in the lower ROV of apprentices.

The inconclusiveness of such labouring was also likely a System 2 phenomenon, which Kahneman defines as 'a metacognitive appreciation of one's ability to think incompatible thoughts about the same thing.'[582] This is an apt description of the observed laboured reasoning and inconclusive assessments of apprentices. It also suggests why masters were able to provide correct and conclusive assessments: they relied more heavily on their System 1 intuition and avoided the increased doubt inherent in System 2 deliberations.

Journeymen's assessments were either incorrect or superficial. This may be explained by their dependency on System 2 reasoning, which either failed to correct their incorrect intuitions and assumptions, or permitted the 'blurting out of whatever comes to mind'[583] because of the mentally demanding nature of the set task. In this latter regard, journeymen were the only participants from groups A and C to assign overriding significance to the fact that the merger parties in Case C had not sought prior approval form the ACCC. Yet as one participant, who provided this kind of assessment, admitted:

> You also would assume that the parties wouldn't have dared completed the acquisition and then inform the ACCC at a later date, if there was anything near a competition problem that would create risk for them … I'll now read on and address the background and go through the factors … but I can see my conclusion already that that alone is going to suggest there are strong prospects that the Commission will not have concerns with this deal[584]

Because this response was reasoned, it was outside the scope of System 1 intuition and is best viewed as superficial System 2 thinking, a fact that this participant seems to have conceded by acknowledging the need to continue with his more detailed assessment. Such behaviour was seen in three other instances where significance was assigned to the post-merger status of this case but the participant still proceeded to

[582] Kahneman, above n 568.
[583] Ibid 699. See also D T Gilbert, 'Thinking Lightly About Others: Automatic Components of the Social Inference Process' in J Uleman and J A Bargh (eds), *Unintended Thought* (Prentice-Hall, 1989) 189.
[584] S17C Line 41-48.

engage in further System 2 reasoning.[585] This contrasts with the master-level participant who immediately and intuitively focused on this issue (in a System 1 sense) and identified it as determinative in a substantive sense.[586] It also contrasts with the other master-level participants who assigned minimal significance to this issue in favour of their more substantive analysis of the case.[587]

In this sense, those participants who provided a superficial but correct assessment in Case C were probably relieved to have identified this non-substantive inference as a short-cut to otherwise laboured and inconclusive reasoning. But those master-level participants who noted this feature were more likely to be focusing their System 2 reasoning on monitoring their intuitive responses to the case, something that their lower-ranked colleagues lacked. The one master who immediately understood the importance of this issue in the context of precedent enforcement action by the ACCC (that is, its failure to secure court orders against a completed merger during the preceding 25 years), relied on intuition confirmed by an encyclopaedic knowledge of this area of law and a highly-efficient indexing of that knowledge within his mental schema.

Other evidence that higher-level participants relied more on their intuition included: the speed at which they identified relevant issues (masters were the fastest); their readiness to rely on their intuition rather than over-think their initial conclusions thereby introducing doubt and inconclusiveness; and their higher ROV when considering cases which is consistent with greater reliance on System 1 rather than System 2 thinking. The fact that some masters were able to quickly provide complete and correct responses to the set task also indicates that they were able to complete the required cognitive tasks concurrently rather than serially, which Kahneman considers 'the most useful indication of whether a given mental process belongs to System 1 or System 2.'[588]

Overall, masters and experts expended less effort completing the test cases because they relied more on their intuition, while the laboured and inconclusive reasoning of apprentices could be attributed to their attempting to reason their way through the cases

[585] S12C, S16C and S19C.
[586] S06C.
[587] S04C, S08C and S05C.
[588] Kahneman, above n 568, 698.

in the absence of intuitive insights. The less laboured identification of issues by journeymen was explicable by their relying on System 2 reasoning to selectively identify issues, but then not fully engaging their System 2 self-monitoring which resulted in incorrect and superficial assessments.

2 *System 2 Self-Monitoring*

As noted in the reasoning options listed by Kahneman above, all judgments involve System 2 reasoning. Part of this is the slow, serial and effortful processes of logical reasoning that is not initiated by System 1 intuition, but rather is a substitute for it (Option 2). Another part of System 2 reasoning involves monitoring System 1 cognition and adjusting the ultimate judgment or decision according to what an individual deems to be the correct or desired outcome. Having concluded that higher-level participants, and in particular master-level participants, in this study likely relied more on System 1 thinking than System 2 reasoning when compared to lower level participants (most obviously apprentices), the role of System 2 in shaping the legal-risk assessments provided by participants is discussed below.

Intuition is considered the main reason why masters and experts, and to a degree journeymen, spent less time identifying issues and retrieving information from their LTM. This explains why these participants spent less time in Stage 2 problem-solving compared to apprentices, with masters taking the least time of all because they relied more on their intuition – and did so more effectively – than any other participants. The relevant data of time spent on this cognitive task according to participants' levels of expertise were shown in Chart 6.1.

Yet in relative terms, experts and masters spent more of their total assessment time drawing inferences and engaging in synthesis. This was shown in Chart 6.2. The question therefore arises as to why they spent more time engaged in self-monitoring when their intuitive responses were ultimately proven correct (and were not overridden by their System 2 reasoning) and their inconclusive assessments were nevertheless unlaboured. More time spent engaging System 2 thinking suggests that there should have been more extensive adjustments to their initial views, more doubt reflected in inconclusive assessments and more laboured thought reflected in lower ROV.

To begin with, equating more time spent in Stage 3 problem-solving (synthesis) with more System 2 thinking is simplistic. System 2 reasoning occurs constantly throughout the legal-risk assessment process, including during times when participants retrieve information from LTM. Assessing whether such information is or is not relevant in relation to a specific legal issue will inevitably involve System 2 analysis.

Second, it is possible for one participant to have engaged in more System 2 reasoning than another participant, but to have actually spent less time engaging in the self-monitoring functions associated with such reasoning. For instance, under Option 2 of Kahneman's model of reasoning note above, there is no System 1 intuition to adjust for. This means that any such adjustments would only apply to inferences and conclusions generated via System 2 thinking. Accordingly, when time is limited and the generation of reasoning in the absence of intuition is effortful, the actual time spent engaged in effective self-monitoring may be minimal even though, overall, System 2 thinking was the predominant reasoning activity.

Third, some inferences and conclusions are likely to require more self-monitoring by virtue of their greater extent and sophistication. Accordingly, several minor or obviously flawed inferences could take significantly less time to check and adjust for than one extended and more sophisticated chain of thought. In this context, the higher the quality of an inference or possible argument, the more time it takes to assess it. Conversely, lower quality inferences and arguments can be quickly dismissed.

Fourth, as noted previously, it is possible to engage in System 2 self-monitoring at varying levels of substantiality. A superficial check for inconsistencies and flaws may not be fully effective, but it will take less time than a more thorough consideration of all relevant issues.

With these factors in mind, it can be hypothesised that apprentices spent less time engaged in self-monitoring than experts and masters, even though they spent more time engaged in System 2 reasoning overall. This is because they were more reliant on such thinking in the absence of reliable intuition. Moreover, their System 2 thinking was focused on identifying issues and making sense of the test cases such that most of the time they spent on this slow type of thinking was included in their spending more time in Stage 2 of the problem-solving process. This left them with less time to self-monitor and suggests that the issues they self-monitored were not as substantial or as high

quality as those generated by higher-ranked participants who relied on intuitive insights. Accordingly, less time was spent and less time was needed for self-monitoring by these lower-level participants.

Journeymen had fewer issues to consider based on their spending less time identifying issues compared to apprentices. However, they spent a roughly equivalent proportion of their total assessment time engaging in synthesis, which on its face suggests that they spent the same proportion of time engaged in self-monitoring – or even more given that they were likely to have generated and then considered fewer substantive issues.

A plausible explanation is that journeymen also spent considerable time generating issues with System 2 thinking while retrieving relevant information from their LTM (Stage 2 problem-solving). Judging from their lower ROV while engaging in this cognitive task, they ostensibly worked harder than apprentices at this task. Insofar as this created more pressure on them to identify relevant issues, they responded by lowering their standards of self-monitoring in an effort to provide a conclusive assessment of some kind, even if such assessments were incorrect or superficial. It is also possible that they were more aware than apprentices of the difficulties of the set task in the absence of intuitive insight, and sought to remedy this by grasping at low-hanging fruit. This could explain why journeymen were attracted to the superficial assessment of Case C (the case involving a completed merger), whereas experts exercised greater control over the substantive quality of their assessments and masters were able to rely more fully on their intuition.

Experts and masters spent, in proportional terms, equal amounts of their assessment time on Stage 3 synthesis. For experts, who spent the same amount of their total assessment time identifying issues as journeymen, they likely spent more time monitoring their responses to the legal-risk assessment task. Even if they spent an equal amount of their System 2 thinking identifying issues, their greater time spent drawing inferences suggests that they would have spent materially more time assessing the quality of those inferences. Insofar as they did not experience the same levels of intuitive insight as masters, experts would have spent more time confirming when they did not have sufficient information on which to form a concluded view. This ostensibly explains why they avoided the incorrect and superficial assessments of journeymen.

Masters arguably spent more time than lower-level participants engaging in self-monitoring as part of their System 2 thinking. They did not need to generate issues using this type of thinking because they relied predominantly on their intuition for this task. Nevertheless, their increased self-monitoring did not result in any major adjustments to their intuitive responses, and there was no evidence that any System 2 adjustments reduced the accuracy or effectiveness of their prior indicated responses. This result accords with Raab and Johnson[589] and Gigerenza and Brighton's[590] contention that such intuitive judgments (to use Kahneman's terminology for System 1 decisions that remain unchanged by System 2 deliberation) conform with a 'take-the-first' heuristic which mediates 'superior expert decision making in highly time-constrained domains.'[591]

Masters also did not experience greater doubt after engaging in more System 2 thinking. This was presumably because of the quality of the issues they identified and because of their ability to avoid focusing on test-case details at the expense of a broader problem-perspective, an attribute that Dijkstra et al suggest protects against the potential detrimental effects of reasoning on judgment.[592]

In these regards, the performance of the masters in this study appears to have confirmed the contentions of Klein,[593] Kahneman[594] and Dijkstra et al[595] that intuition rather than deliberation in the kinds of information-limited and time-constrained situations used in this study is likely to yield the best results. It also indicates that masters have developed a greater trust in their own intuition and that they resist making adjustments

---

[589] M Raab and J G Johnson, 'Expertise-Based Differences in Search and Option-Generation Strategies' (2007) 13 *Journal of Experimental Psychology: Applied* 158.
[590] G Gigerenzer and H Brighton, 'Homo Heuristicus: Why Biased Minds Make Better Inferences' (2009) 1 *Topics in Cognitive Science* 107.
[591] Jerad H Moxely, K Anders Ericsson, Neil Charness and Ralf T Krampe, 'The Role of Intuition and Deliberative Thinking in Experts' Superior Tactical Decision-Making' (2012) 124 *Cognition* 72, 73. However, this contention is not supported by Moxley et al's findings that in chess all levels of experts and non-experts benefit from deliberation when faced with tactical problems, although this appears to depend (and the benefit alternates) based on the level of difficulty of the set task and on the fact that Moxley et al only considered single tactical-chess moves rather than complete games. Experts in this study of chess expertise benefited more than non-experts from deliberation in difficult cases. In easy cases, non-experts benefitted more. Ibid 76.
[592] 'Deliberation induces a local processing style in which people tend to focus on details and pay less attention to the global picture. This focus on details can explain the detrimental effects on judgment.' Dijkstra, van der Pligt and van Kleef, above n 576, 286.
[593] Klein, above n 568.
[594] Kahneman, above n 568.
[595] Dijkstra, van der Pligt and van Kleef, above n 576.

to their intuitive insights, even when new information (like the low-hanging fruit grasped by journeymen in Case C) appears.[596]

This analysis is also supported by Klein's conclusions[597] that highly expert decision-makers do not choose between options, but rather focus on only one or two, running each to their ultimate conclusion through a rigorous self-monitoring process. This was likely reflected in the master-level participants in this study spending less time identifying issues and more time drawing inferences and making sure everything was in alignment to provide a high-quality and correct assessment of legal risk.

3 *Response to the Research Question*

This response to the research question offers an explanation for the readily identifiable and measurable differences noted above in parts A, B and C of this chapter. It does not identify any further differences of that kind, although it may be possible in some circumstances to do so using these findings either directly or indirectly. The better view is that the present analysis is of a complementary nature insofar as it provides further depth to the analysis undertaken in the previous three parts of this chapter.

Apprentices in this study relied heavily on System 2 thinking to identify issues and retrieve information from their LTM. This approach was effortful and led to laboured and inconclusive assessments. Apprentices' extensive use of System 2 reasoning, which involved detailed analyses of conflicting considerations, increased doubt and indecision. These participants engaged in limited self-monitoring of a relatively large number of low-quality issues. A key distinguishing feature of apprentices thinking was their overconfident reliance on their intelligence (reasoning) over their experience (intuition), which resulted in laboured inconclusiveness when assessing legal risk under the test conditions of this study.

Journeymen also relied on System 2 reasoning to identify issues, but did not identify as many as apprentices. It is possible that they sensed the futility of a purely reasoned response, perhaps because they were more globally aware. However, in the absence of

[596] This tunnel vision is likely to be explained at least in part by Kahneman's findings that intuitive thinking is less flexible than deliberative thought. Kahneman, above n 568, 698.

[597] See Klein, above n 568; Klein, above n 576; G A Klein, 'A Recognition Primed Decision (RPD) Model of Rapid Decision Making,' in G A Klein, J Orasanu, R Calderwood and C E Zsambok (eds), *Decision-Making in Action: Models and Methods* (Ablex, 1993) 138; G Klein, *Sources of Power: How People Make Decisions* (MIT Press, 1998); Gary Klein, 'Naturalistic Decision Making' (2008) 50(3) *Human Factors* 456.

major intuitive insights (and under the pressure of the testing process) they searched for low-hanging fruit as a means of providing at least some form of concluded view, even when this led to superficial or incorrect results. These participants did not undertake extensive self-monitoring but spent less time on this task further contributing to their low quality assessments. The behaviour of blurting out responses to the legal-risk assessment task ostensibly reflected a response to the constraints of the test procedures and inadequate self-monitoring. It may have also reflected journeymen's attempted substitution of guessing for intuition.

Experts maintained a balanced reliance on System 1 and System 2 thinking. However, insofar as they did not have any significant intuitive insights they could not provide conclusive assessments of legal risk. Nevertheless, their self-monitoring was extensive and they avoided the low-quality responses that journeymen provided. They also considered fewer issues than journeymen, which allowed them to spend more of their time monitoring their System 2 reasoning. Experts refused to be drawn into guessing the likely outcomes of the test cases, and when they determined that they had insufficient information they consciously avoided giving a response. This, in addition to an absence of laboured reasoning, indicated effective self-monitoring within their System 2 processing.

Masters relied most on their System 1 intuition and were, as a consequence, focused on just one or two key issues. This reduced the time they spent retrieving information from LTM and allowed them more time to engage in extensive self-monitoring. This self-monitoring was effective insofar as it neither adversely affected their correct intuitions nor introduced doubt into their assessments. In this latter respect, they avoided the apprentice's mistake of struggling with an overly detailed focus on the facts of the test cases at the expense of the bigger picture. They therefore escaped the tendency of deliberative thought to focus attention only 'on accessible and reportable information.'[598] Masters' greater experience both increased their chances of correct intuitive responses to the given task and ensured that these participants would be

[598] Dijkstra, van der Pligt and van Kleef, above n 576, 291. See further: R S Tordesillas and S Chaiken, 'Thinking Too Much of Too Little? The Effects of Introspection on the Decision-Making Process' (1999) 25 *Personality and Social Psychology Bulletin* 623; T D Wilson, S D Hodges and S J LaFleur, 'Effects of Introspecting About Reasons: Inferring Attitudes from Accessible Thoughts' (1995) 69 *Journal of Personality and Social Psychology* 16.

willing and able to capitalise on such responses without overthinking their deliberations.

E *Other Factors*

The above discussion focused on identifying, measuring and explaining the different legal-risk assessment behaviours of study participants with different levels of specialist legal expertise. These behaviours covered the ease with which participants undertook the legal-risk assessment task (laboured vs unlaboured), the certainty of their assessments (conclusive vs inconclusive), the accuracy of their assessments (correct vs incorrect – conclusive assessments only), and the depth of their analysis (superficial vs substantive – conclusive and correct assessments only). In each of these areas, quantitative data – combined with problem-solving theory and findings from research on decision-making and judgment – were used to explain the observed differences between higher-ranked legal specialists and lower-ranked legal specialists.

To facilitate this analysis, special care was taken to select study participants who shared common professional backgrounds and specialist technical knowledge. Nevertheless, it was not possible to exclude all factors that could have impacted on and influenced the observed data in ways not fully captured by the results and chosen theoretical approach.

Below is a brief discussion of the more obvious of these factors and their possible effects on the above analysis.

1 *External Time Pressures*

All prospective participants were asked to set aside 45 minutes to one hour of their time to be interviewed for this study. The fact that no individual sought to end an interview in less time suggests that no participant was under pressure to rush their assessment of any test case, which were nominally presented as requiring 10-minutes each to assess.

However, it is conceivable that some participants – perhaps more likely the more junior ones – were concerned about being seen within their organisations as wasting time on an activity for which neither they nor their firm would benefit directly. Moreover, these participants may have been more susceptible to (and more aware of the possibility of) being abruptly interrupted and asked to work on other things immediately. The more senior participants were perhaps less likely to have had this concern.

One behaviour potentially affected by this influence was correct conclusions based on superficial analyses. External time pressures could have provided the motivation to select the fastest method for assessing legal risk, which for some participants may have resulted in a tendency towards guessing, analytical short-cuts and perfunctory self-monitoring. Their aim could have been simply to get through the cases as quickly as possible, by whatever means.

While this possibility cannot be ruled out, it seems unlikely it was something that had a disproportionate influence on more junior participants, that is, those participants in Group C categorised as apprentices and journeymen. In addition to the three Group C participants who relied on superficial analyses in forming their opinions, there were three Group B participants who adopted the same strategy. The first of these latter participants had 13 years professional experience, the second 23 years professional experience, and the third 35 years professional experience.

More generally, the fact that all participants had volunteered their time and were not pressured to do so suggests that they were favourably disposed to the aims of the study and were motivated to provide their fullest involvement. Insofar as each participant had freedom to schedule their interviews for convenient times and always had the option of re-scheduling them, concerns that they were subject to the more obvious pressures of their workplace were allayed.

2 *Internet Searching*

Participants were required to follow prompts, answer questions and read test-case documentation on the study's testing website. Being connected to this on-line portal in their offices meant that they would also have had access to Internet search engines and other resources to help them complete the assessment task. For instance, they could have searched for information on merger parties or found out more about relevant products or details of the industries concerned. This potential might have favoured younger participants in Group C more than Group A participants, who may have been less likely to automatically undertake such searches as questions arose in their minds.

This possibility cannot be ruled out. However, it seems improbable that on-line searches were undertaken given the fulsome nature of all participants' verbalisations and the fact that many participants – including those in Group C – completed

assessment tasks while admitting that they did not fully understand the nature of the industry or the relevant products or services. These factors suggested that Internet searching was not an actual issue of concern, even if the possibility existed by virtue of the study's reliance on on-line testing.

3 *Novelty of the Testing Procedure*

It was clear from the responses of all participants that they had never been asked to undertake legal-risk assessments in the manner required for this study. At the same time, no participant indicated that they did not understand what their task was. Indeed, most appeared able to immerse themselves quickly into the role of adviser in each case, with some commenting that they had had experienced analogous situations during their professional careers, although not in this precise format.

It is nevertheless possible that some more senior participants could have had greater difficulty adapting to the on-line format of the testing process and more generally felt unused to having such strict limitations on available information and time. As noted in the previous chapter in relation to those participants who refused to provide an assessment in some cases, these were senior practitioners who may have been more dependent on having greater amounts of contextual information than was made available to them.

Apart from this possibility, there was no actual indication that Group A and Group C participants reacted differently to the novelty of the test procedures in terms of its hypothetical setting or reliance on restricted information in the form of on-line documentation.

4 *Mergers Experience*

While the majority of participants indicated that merger review cases were a specific focus of their professional practices, the number of such matter on which they had worked during the prior 12 months, 5 years and their careers-to-date varied. However, when comparing Group A participants and Group C participants, the only statistically significant difference between them in this area was that Group A participants, as a whole and on average, had been involved in more merger matters than Group C participants during their careers-to-date. This was consistent with their having had longer careers.

It is also important to recall that merger cases were used as test cases in this study because they raised a single, substantive question of law ('Is this merger likely to substantially lessen competition?') and this question required the analysis of issues common not just to merger review cases, but also to a wide range of other cases commonly handled by competition law specialists. One of the most important and generic of these issues was the delineation of relevant product and geographic markets, which is central to most assessments of legal risk in competition law cases not just merger cases.

In terms of ease of reasoning, those Group C participants who engaged in laboured reasoning were generally less experienced in merger matters, although one of these participants had been involved in 50 such matters during the last five years, which was the third highest response across all participants for this time period. In addition, the three Group B participants who engaged in laboured reasoning at least twice had been involved in 30, 35 and 100 matters during their careers-to-date. Moreover, one of the highest ranked and best performing participants in Group A had worked on only 10 merger cases during his 20-year career.

There was a similar result when comparing those Group C participants who provided inconclusive assessments or incorrect opinions with those participants in groups A and B who performed better but who had similar (and in some instance less) experience in terms of their involvement in previous merger matters. As noted in Section 1 above, correct assessments based on superficial analyses were not restricted only to more junior participants, nor to participants who had worked on only a limited number of merger matters. Two of the Group B participants who completed such assessments had each worked on over 100 merger cases during their careers.

However, in terms of incorrect assessments considered as stand-alone outcomes, there appeared to be a possible association between how many merger matters these participants had considered during their careers to date and their incorrect conclusions. Indeed, this may have been a reason why the sole Group A participant, who had only been involved in 10 merger matters during his 20-years career, provided an incorrect assessment. The other four Group C participants who had provided incorrect assessments had career statistics of 6, 5, 20 and 50 merger matters. However, the one

Group B participant who provided an incorrect assessment had been involved in more than 150 merger matters during his career.

An additional complicating factor was the difficulty of determining what constituted involvement in a merger matter. In one instance, it may have involved primary carriage of major transaction all the way through the merger review process. In another, it could have been half a day's research to determine that no merger clearance application was required.

Lastly, with respect to the Group A participant who incorrectly believed that the transaction in Case B would be opposed (and who, as noted, had been involved in only 20 merger matters during his career), he provided a correct assessment in Case C based on a substantive analysis of the key issues in that case. No Group C participant who provided an incorrect assessment also provided a correct assessment on such a basis in any case. It is also recalled that this Group A participant's performance in Case B was more similar to other Group A assessments than Group C assessments on each of the quantitative measures discussed above.

5 *Professional Focus*

All participants were asked whether their specialist legal practices were focused more on front-end advisory work, back-end litigation work, or a balance of both front-end and back-end work. There was no statistically significant difference between the mix of practices amongst Group A participants and amongst Group C participants.

In relation to laboured reasoning, two Group C participants who engaged in this form of reasoning focused on back-end work, one focused on front-end work, and two had balanced practices. Amongst the three Group B participants who engaged in laboured reasoning, one focused on back-end work, one on front-end work and one had a balanced-practice. Laboured reasoning therefore did not appear to be related to a participant's professional focus.

There was a similar mix of professional practices amongst those participants who provided inconclusive assessments, while the six participants who provided incorrect conclusive assessments had back-end (two participants), front-end (one participant), and balanced (three participants) practices. The practices of the 14 participants who provided correct conclusive assessments were similarly varied.

In terms of the three Group C participants who based their correct assessments on superficial analyses, each had a different practice focus. Similarly, amongst the four Group B participants who provided this kind of assessment, one had a balanced practice, one a front-end practice, and two had back-end practices. This same variation was evident amongst the five Group A participants who provided correct assessments (Two had front-end practices, two had balanced practices, and one had a back-end practice).[599]

6 *Economics Qualifications*

Given that the assessment of legal risk in competition law cases typically involves economic analysis – as was the case in each test-case in this study – having economics qualifications could have given some participants an advantage in completing their assessments of merger cases. There were six participants who had economics qualifications and who also practiced as competition economists. Another two had economics qualifications but practiced as lawyers.

As between groups A and C, there was no statistically significant greater incidence of practising competition economists in one group or the other. However, one participant in Group C was a lawyer with economics qualifications. None of the lawyers in Group A had economics qualifications.

The five Group C participants who engaged in laboured reasoning were lawyers, including the one who also had economics qualifications. In Group B, two of the three participants who engaged in laboured reasoning were economists, and the third was a lawyer without economics qualifications. Unlaboured reasoning was engaged in by economists and lawyers with similar levels of incidence in each category and across all groups.

Conclusive and inconclusive assessments were provided by both economists and lawyers, and lawyers with economics qualifications. Correct and incorrect assessments and correct assessments based on superficial analyses and on substantive analyses were similarly provided by a mixture of economists and lawyers.

[599] Amongst the three Group A participants who provided two or more correct assessments, one had a front-end practice, one had a balanced-practice, and the last had a back-end practice.

7 *Degree of Specialisation*

All participants were asked to state the degree to which they had specialised in competition law during the previous 12 months. This calculation was based on how much of their billable time during that period was spent on competition law matters.

Amongst the 11 participants who had spent more than 50% but less than 75% of their time during the last 12 months on competition law matters, three had engaged in laboured reasoning , two had provided three or more inconclusive assessments, five had given incorrect assessments, and eight had given correct assessments, of which three were based on superficial analyses.

Amongst the nine participants who had spent 75% or more of their time during the last 12 months on competition law matters, two had engaged in laboured reasoning, four had provided three or more inconclusive assessments, one had given an incorrect assessment, and six had given correct assessments, of which three were based on superficial analyses.

The two areas of potential difference between these statistics concerned inconclusive assessments and incorrect assessments. Generally, those participants who had specialised for more than 75% of their time were twice as likely to provide inconclusive assessments, while those who had specialised less than 75% of their time were four times more likely to provide incorrect assessments.

Why those participants who specialised more should provide more inconclusive assessments is unclear. A possible explanation is that this variable was overshadowed by the fact that only some Group B and Group C participants (but no Group A participants) provided three or more inconclusive assessments, suggesting that it was their level of expertise rather than their degree of specialisation that contributed to their records of fewer conclusive assessments.

Why those participants who specialised less provided more incorrect assessments may be explained by their lacking current knowledge of recent merger review cases and therefore being less familiar with the ACCC's current views on specific issues. However, simply specialising less in competition law does not necessarily mean involvement in fewer merger matters compared to someone else who spends all their

time in this field of law, but who has been involved in one large litigation case for two years.

Moreover, as discussed next, three of the six participants who provided incorrect assessments had previously worked as a member of staff at the ACCC, with one having worked specifically in the ACCC's mergers branch.

8 *Regulatory Experience*

Ten participants in this study had either prior or current experience working as a member of staff at a competition authority. In respect to laboured reasoning, incorrect assessments, correct assessments based on substantive analyses and correct assessments based on superficial analyses, there were no material differences between these participants, taken as a group, and those participants without experience working with a competition authority.

In respect to inconclusive assessments, however, those participants who had worked or were working with a competition authority were twice as likely to provide such an assessment as compared to other participants. However, it is possible that this was not a significant finding inasmuch as it relates to only six participants, four of whom had competition authority experience and two of whom did not.

F *Summary and Conclusion*

The following table summarises the analysis in parts A, B, C and D of this chapter. It maps the key differences identified between the different levels of legal specialists who participated in the information-constrained and time-limited legal-risk assessment tests conducted during this study. For each level of legal specialist, the key behavioural identifiers, performance traits and cognitive indicators are shown, with those indicators unique to a particular level of expertise in bolded text. Participants' expertise-related reliance on intuition and analytical reasoning are presented as a sub-category of their performance traits as shown in italics in Column 2.

TABLE 6.1 – Summary of Expertise Identifiers, Performance Traits and Cognitive Indicators

	1. IDENTIFIABLE BY	2. PERFORMANCE TRAITS	3. COGNITIVE INDICATORS
MASTER	• **Correct legal-risk assessments based on substantive analyses** • Unlaboured reasoning • [Refusals to give an assessment when insufficient information available][600]	• **Identifies key issues easily, and very quickly** • Avoids irrelevant and low-quality issues • Uses appropriate and specific analogies • Ignores/glosses-over short-cuts that minimise synthesis • Synthesises issues efficiently and effectively • *Heavy and effective reliance on intuition* • *Extensive and effective self-monitoring* • *Avoids deliberative doubt*	• **Very low proportion of total assessment time spent retrieving information from LTM** • High proportion of total assessment time spent drawing inferences • Low overall cognitive effort • Consistently high ROV whether retrieving information from LTM or drawing inferences
EXPERT	• Unlaboured reasoning • Inconclusive assessments • Refusals to give an assessment when insufficient information available	• **Identifies key issues easily** • Avoids irrelevant and low-quality issues • Uses appropriate and specific analogies • Ignores/glosses-over short-cuts that minimise synthesis • Synthesises issues efficiently and effectively • *Integrated intuition and reasoning* • *Extensive and effective self-monitoring* • *Limits deliberative doubt*	• Low proportion of total assessment time spent retrieving information from LTM • High proportion of total assessment time spent drawing inferences • Low overall cognitive effort • Consistently high ROV whether retrieving information from LTM or drawing inferences
JOURNEYMAN	• **Correct legal-risk assessments based on superficial analyses** • Unlaboured reasoning • Inconclusive assessments • Incorrect legal-risk assessments	• **Issue identification is laboured**. • **Identifies short-cuts that minimise synthesis** • Identifies irrelevant and low-quality issues • Synthesis is inefficient and limited/superficial • *Substitutes guessing for intuition* • *Limited and ineffective self-monitoring* • *Tendency to blurt out premature responses*	• Low proportion of total assessment time spent retrieving information from LTM • Low proportion of total assessment time spent drawing inferences • High overall cognitive effort • Significantly lower ROV when retrieving information from LTM than when drawing inferences
APPRENTICE	• **Laboured reasoning** • **Inconclusive assessments** • Incorrect legal-risk assessments	• **Issue identification is laboured, and extensive** • **Uses vague and inappropriate analogies** • Identifies irrelevant and low-quality issues • Synthesis is inefficient and limited/superficial • *Heavy reliance on analytical reasoning* • *Limited self-monitoring* • *Overwhelmed by deliberative doubt*	• **High proportion of total assessment time spent retrieving information from LTM** • Low proportion of total assessment time spent drawing inferences • High overall cognitive effort • Significantly lower ROV when retrieving information from LTM than when drawing inferences

Each of the differences noted in this table is supported by the observations and statistically significant quantitative comparisons recorded in Chapter 5, and by the foregoing analysis. For the reasons previously discussed, these differences do not appear attributable to the factors and potential confounding influences identified in Part E. To the extent permitted by the limitations of a study of this kind, the selection of participants and the testing methodology used in this study were focused on – and were

[600] Insofar as master-level participants may have refused to provide an assessment in certain cases because of lack of information provided in the test-case documentation – which responses were arbitrarily categorised as expert-level only – mention is made of this as identifiable assessment behaviour amongst masters as well as amongst experts.

assumed to be largely effective in – highlighting cognitive differences associated with different levels of specialist legal expertise.

Not all the differences that might exist in this area were recorded in this study. The sample-size was not large, although it was larger than previous studies of a similar type. The representative task chosen to test the cognitive abilities of study participants was of a specific kind and could not encapsulate all possible activities that require specialist legal skills or which could vary depending on the level of expertise of the individuals involved. Further, not all the data generated in this study were or could be subjected to analysis – only those data identified and analysed in this chapter.

Bearing these considerations in mind, the above table offers a useful summary of the findings of this study. More particularly, it provides an overview of the key issues relevant to answering the research question posed at the beginning of this thesis, namely:

> In what readily identifiable and measurable ways do legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts?

Considering in turn each level of expertise according to Hoffman's Scheme and the traditional model of progressive expertise development, the following sections summarise the distinguishing features of the different cognitive approaches adopted by the participants in this study as they undertook the representative legal-risk assessment task. The analysis undertaken in Part D, which focused on the use of intuitive and analytical reasoning, is also included in these summaries. While not strictly required as a response to the research question, this additional analysis offers an explanatory context in which to assess the above-listed measures.

1 *Apprentices*

The apprentices in this study were identifiable by their laboured reasoning and inconclusive assessments. Laboured reasoning was not identified amongst experts and masters. Inconclusive assessments were noted amongst higher-level participants, but they were predominantly associated with apprentices.

In terms of identifying issues, apprentices undertook both laboured and extensive searches for relevant information. However, this information was often irrelevant or of low-quality. An example was their poor use of analogies. Given that analogies were an important input into the assessment process because the test-case documentation had very little descriptive information and participants' LTMs were also lacking in direct, relevant information about the industries concerned, an inability to use this cognitive skill effectively disadvantaged lower-level legal specialists.

These performance behaviours were reflected in apprentices spending a high proportion of their total assessment time retrieving information from their LTM. Like journeymen, this cognitive task required greater cognitive effort than they expended on drawing inferences, which were less of a priority and more superficially undertaken.

The key elements of apprentices reasoning strategies in terms of System 1 and System 2 thinking were their: heavy reliance on System 2 analytical reasoning in the absence of reliable System 1 intuitive responses; limited System 2 self-monitoring in terms of effective testing and control of their reasoned responses; and high levels of doubt attributable to excessive deliberations over low-quality details.

2 *Journeymen*

Journeymen were most readily identifiable by their providing correct assessments of legal risk based on superficial analyses. Unlike experts and masters, they appeared to actively identify opportunities to avoid engaging in synthesis. When no such opportunities were available, their conclusions were incorrect, possibly because they did not spend enough time on synthesis. This may have been because like apprentices they had to work hard to identify relevant issues (although unlike apprentices they did not undertake as extensive searches for relevant information). It is also possible that the relatively low-quality of the information retrieved from their LTM was unlikely to yield better results even with more time drawing inferences.

These behaviours were reflected in or were associated with journeymen's unique combination of spending a low proportion of their time both on retrieving information from LTM and drawing inferences. They spent 30% more time, in a proportional sense, reading and clarifying the test case documentation than apprentices did. This may have been because they were more focused on searching the provided data for

opportunities to undertake superficial analyses in the absence of high-quality retrieved information and their ability to synthesise such information as efficiently and effectively as higher-level experts.

Like apprentices, journeymen also relied on System 2 reasoning to identify issues, but did not identify as many as apprentices did. Their limited self-monitoring was insufficient to prevent the blurting out of premature responses. It also permitted the superficial guessing of issues and outcomes in the absence of intuitive insights.

3 *Experts*

Experts were similar to journeymen in that they spent less of their total assessment time identifying issues compared to apprentices. But unlike journeymen, experts avoided identifying irrelevant and low-quality issues. They were also better able to synthesise those issues. Perhaps for this reason they spent a much higher proportion of their assessment time drawing inferences and maintained a consistent level of cognitive effort whether retrieving information from LTM or drawing inferences. This combination of less time identifying issues and more time synthesising was unique to experts.

Some experts in this study refused to provide an assessment because they had insufficient information to do so conclusively. This may have reflected their greater dependence on contextual information. They also glossed-over apparent analytical short-cuts as did masters, and were able to identify key issues quickly, although not as quickly as masters.

Experts demonstrated a more moderated reliance on System 2 thinking. Like journeymen they lacked determinative intuitive insights. However, their self-monitoring was extensive, although not excessive, and they avoided the low-quality responses that journeymen provided to the point of refusing to provide any assessments of legal risk.

4 *Masters*

Masters in this study were, by definition, the only participants capable of providing correct assessments of risk based on substantive legal and economic analyses. They were also the fastest at identifying key issues, which explained why they spent the least

amount of their total assessment time retrieving relevant information from their LTM. They otherwise spent essentially the same proportion of their time drawing inferences as experts did. This combination of very little assessment time spent identifying issues and a large proportion of time engaging in synthesis was a distinguishing characteristic of masters.

The insufficiency of provided information may have led some masters to refuse to provide any opinion on legal risk and thus their assessments were characterised as expert-level only. They glossed-over analytical shortcuts, avoided irrelevant and low quality issues, demonstrated low overall cognitive effort compared to journeymen and apprentices, and their synthesis of issues required no more cognitive effort than identifying issues but was demonstrably efficient and effective.

Masters were the most intuitive of all participants inasmuch as their System 1 intuition was both reliable and correct. Their System 2 reasoning was also effective insofar as it neither adversely affected their correct intuitions nor introduced excessive doubt into their assessments. On this basis, they avoided getting bogged down in factual details and did not overthink their assessments of legal risk.

5 *Conclusion*

The results of this study as recorded in the previous chapter and analysed above, form the basis for a list of readily identifiable and measurable ways in which legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts. This list comprises the qualitative identifiers in Column 1 of Table 6.1 and the quantitative cognitive indicators in Column 3 of the same table. Column 2 of this table provides explanatory information relating to the problem-solving and decision-making performances of the different levels of legal specialists in this study, although these are largely based on inference and are therefore less easily identified and measured directly.

These differences were summarised and compiled according to the categories of apprentice, journeyman, expert and master. These categories relied on the conceptual framework of Hoffman's Scheme which was derived from the developmental ranking of expertise used by Middle Ages craft guilds. The particularisation and specificity of

these category summaries provides for the first time a basis for distinguishing between individual legal specialists with different levels of expertise in the same field of law.

The next chapter concludes this thesis with a review of the studies and research on which this study relied. It also provides a consideration of the implications of the above findings for previous scholarship concerned with the cognitive analysis of legal expertise and for future research opportunities.

VII CONCLUSION

This thesis described the methodology and results of an empirically-based response to the following research question:

> In what readily identifiable and measurable ways do legal specialists with different levels of expertise (but the same technical legal knowledge) think differently when assessing legal risk in information-limited and time-constrained contexts?

Part F of the previous chapter lists a number of qualitative identifiers, performance traits and cognitive indicators that provide a means of distinguishing four levels of legal specialists according to their levels of expertise rather than their legal knowledge. This is the first time such a list has been constructed. Based on the findings of this study, a number of further research opportunities now exist with the potential for advances in relation to: the cognitive development of legal expertise; the formal accreditation of legal specialists; assessments by consumers of legal services; the development of legal talent within law firms; and legal education, particularly at intermediate and advanced levels.

This chapter is divided into four parts.

Part A reviews the previous research on which this thesis relied. It then describes areas where the findings of this thesis appear to confirm the findings of previous studies regarding the cognitive abilities of legal experts. It also highlights a number of areas where, in light of the results of this study, a reassessment of previous studies may be warranted. This includes a discussion of the similarities and differences between this thesis and previous research in terms of both methodology and results.

Part B discusses limitations affecting and resulting from the scope and investigatory approach of this thesis. These include the implications of using verbal protocol analysis to assign participants to their initial expertise groups, the exploratory and map-like analytical approach of the study, the focus on those participants initially ranked above and below the main body of participants, the restricted utilisation of these un-ranked participants, the decision to not employ additional assessors, the inherent inefficiencies of the chosen analytical approach, and the missed opportunity to require

participants to provide a percentage estimate as to the likely success of each case they considered.

Part C identifies specific areas of future research which may be facilitated by or benefit from the manner in which this thesis was developed and the response it has provided to the above research question. These areas of future research extend across the disciplines of cognitive psychology, lawyer accreditation schemes, consumer information, law-firm talent management and legal education. In addition, there is the opportunity to replicate this study and reproduce the results of this thesis using the detailed methodological descriptions provided above and the availability of the actual on-line testing platform developed during the course of this research project.

Part D offers a concluding statement to this thesis.

A *Previous Research*

1. *Reliance*

This thesis relied on established sources of theoretical and empirical research from the fields of cognitive psychology and expertise studies.[601] It was more specifically guided by studies that have used think-aloud problem solving and verbal protocol analysis to investigate legal thinking skills.[602] These studies included contributions by researchers

[601] These sources include: K Anders Ericsson and Herbert Simon, *Protocol Analysis: Verbal Reports as Data* (The MIT Press, 1984/93); K Anders Ericsson, 'Protocol Analysis and Expert Thought: Concurrent Verbalizations of Thinking during Experts' Performance on Representative Tasks' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 223; the contributions of medical researchers such as Patel and Groen (V L Patel and G J Groen, 'Developmental Accounts of the Transition from Medical Student to Doctor: Some Problems and Suggestions' (1991) 25(6) *Medical Education* 527); and the pioneering work of the Berlin Group – P B Baltes and U M Staudinger, 'Wisdom: A Metaheuristic (Pragmatic) to Orchestrate Mind and Virtue Toward Excellence' (2000) 55 *American Psychologist* 122; P B Baltes, U M Staudinger, A Maercker and J Smith, 'People Nominated as Wise: A Comparative Study of Wisdom-Related Knowledge' (1995) 10(2) *Psychology and Aging* 155; Staudinger, U M, J Smith and P B Baltes, 'Wisdom-Related Knowledge in a Life Review Task: Age Differences and the Role of Professional Specialization' (1992) 7 *Psychology and Aging* 271; Staudinger, Ursula M, Jacqui Smith and Paul B Baltes, *Manual for the Assessment of Wisdom-Related Knowledge* (Max Planck Institute for Human Development and Education, 1994); J Smith and P B Baltes, 'Wisdom-Related Knowledge: Age/Cohort Differences in Response to Life-Planning Problems,' (1990) 26 *Developmental Psychology* 494.

[602] For example: Crombag et al's failed attempts in the 1970s (H F M Crombag, J L De Wijkerslooth and E H Van Tuyl Van Serooskerken, 'On Solving Legal Problems' (1975-1976) 27 *Journal of Legal Education* 168); the contributions of Lundeberg (Mary A Lundeberg, 'Metacognitive Aspects of Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407), Christensen (Leah M Christensen, 'Legal Reading and Success in Law School: An Empirical Study' (2006-2007) 30 *Seattle University Law Review* 603), Oates (Laurel Currie Oates, 'Leveling the Playing Field: Helping Students Succeed by Helping Them Learn to Read as Expert

such as Nievelstein et al,[603] Colon-Navarro,[604] Weinstein[605] and Chay[606] who focused on the cognitive analysis of specialist legal expertise, albeit primarily within the framework of the knowledge-based, novice-expert dichotomy.

A key study-design decision influenced by these prior studies was the choice of a representative task to test participants' subdomain expertise rather than merely their generic legal skills – or simply their ability to complete an artificial task lacking ecological validity. Direct guidance was taken from those studies that provided theoretical and practical commentary on these and related issues, such as the appropriate instructions to give to participants and how to eliminate case-specific knowledge advantages.

The initial ranking of study participants relied on the cumulative research by a variety of scholars,[607] including Hoffman[608] from whose research Hoffman's Scheme[609] was

---

Lawyers' (2006) 80 *St John's Law Review* 227), Deegan (Dorothy H Deegan, 'Exploring Individual Differences Among Novices Reading in a Specific Domain: The Case of Law' (1995) 30(2) *Reading Research Quarterly* 154) and Stratman (James F Stratman, 'When Law Students Read Cases: Exploring Relations Between Professional Legal Reasoning Roles and Problem Detection' (2002) 34(1) *Discourse Processses* 57) on generic legal skills mainly relating to reading; and the theoretical discussions of Mitchell (John B Mitchell, 'Current Theories on Expert and Novice Thinking: A Full Faculty Considers the Implications for Legal Education' (1989) 39 *Journal of Legal Education* 275), Blasi (Gary L Blasi, 'What Lawyers Know: Lawyering Expertise, Cognitive Science, and the Functions of Theory' (1995) 45(3) *Journal of Legal Education* 313) and Krieger (Stefan H Krieger, 'The Development of Legal Reasoning Skills in Law Students: An Empirical Study' (2006) 56(3) *Journal of Legal Education* 332) regarding the utility of undertaking the cognitive analysis of substantive legal expertise.

[603] Fleurie Nievelstein, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Expertise-Related Differences in Conceptual and Ontological Knowledge in the Legal Domain' (2008) 20(6) *European Journal of Cognitive Psychology* 1043.

[604] Fernando Colon-Navarro, 'Thinking Like a Lawyer: Expert-Novice Differences in Simulated Client Interviews' (1997) 21 *The Journal of the Legal Profession* 107.

[605] Ian Weinstein, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1.

[606] Allan James Chay, *Lawyer Problem Solving: An Investigation of the Knowledge Used in Solving Practical Legal Problems* (PhD Thesis, Griffith University, 2006).

[607] Principally: M T H Chi, 'Two Approaches to the Study of Experts' Characteristics,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 21; M T H Chi, P Feltovitch and R Glaser, 'Categorization and Representation of Physics Problems by Experts and Novices' (1981) 5 *Cognitive Science* 121; Ericsson, above n 601; David Z Hambrick, Frederick L Oswald, Erik M Altmann, Elizabeth J Meinz, Fernand Gobet and Guillermo Campitelli, 'Deliberate Practice: Is That All It Takes To Become An Expert?' (2013) 45 *Intelligence* 34; Phillip L Ackerman, 'Nonsense, Common Sense and Science of Expert Performance: Talent and Individual Differences' (2013) 45 *Intelligence* 6; Lesgold et al (1988); Weinstein, above n 605; Chay, above n 606; Staudinger et al, *Manual for the Assessment of Wisdom-Related Knowledge*, above n 601; C E Greaves, H Zacher, B McKenna and D Rooney, 'Wisdom and Narcissism as Predictors of Transformational Leadership' (2014) 35(4) *Leadership & Organization Development Journal* 335.

[608] Robert R Hoffman, 'How Can Expertise be Defined? Implications of Research from Cognitive Psychology' in Robin Williams, Wendy Faulkner and James Fleck (eds), *Exploring Expertise: Issues and Perspectives* (Macmillan Press Limited, 1998) 81.

[609] For a summary of Hoffman's Scheme see Table 2.2 in Chapter 2.

derived as a means of categorising different levels of expertise in a traditional progressive sense. The chosen testing methodology involving limited information and time-constrained interviews relied on practical guidance from other researchers,[610] theoretical insights from the acknowledged pioneers in the field[611] and examples from those who have focused on lawyer-specific assessments.[612] The exploratory nature of the subsequent analysis relied on the work of still other theoreticians and practitioners.[613]

The main points of departure from prior studies were the classification of novel legal-risk assessment behaviours that were derived from observation, and the generation of quantitative data relating to the participants' focus on different cognitive tasks. The use of verbalisation rates to measure cognitive load was not new, although the division of transcripts into 10-second intervals to facilitate time and task analysis was a variation on Deffner's approach[614] that has not been previously utilised. These were considered appropriate adaptations to facilitate the cataloguing of readily identifiable and measurable differences between the different levels of legal specialists involved in this study.

2 *Confirmation*

This thesis is primarily concerned with investigating new perspectives on how legal experts think with reference to the traditional progressive framework of Hoffman's Scheme, which describes the categorisation of apprentices, journeymen, experts and masters. Further, the emphasis in previous chapters was on finding readily identifiable and measureable cognitive differences, not just theoretical or purely inferential ones.

---

[610] For instance: Hoffman, above n 608; Beth Crandall, Gary Klein and Robert R Hoffman, *Working Minds: A Practitioner's Guide to Cognitive Task Analysis* (The MIT Press, 2006).
[611] Ericsson and Simon, above n 601.
[612] Colon-Navarro, above n 604; Weinstein, above n 605; Chay, above n 606.
[613] For example: A Baddeley, 'Working Memory: Theories, Models, and Controversies' (2012) 63 *Annual Review of Psychology* 1; Gary Klein, 'Naturalistic Decision Making' (2008) 50(3) *Human Factors* 456; Gary Klein, Karol G Ross, Brian M Moon, Devorah E Klein, Robert R Hoffman and Erik Hoffnagel, 'Macrognition' (2003) 3 *Human-Centered Computing* 81; Paul J Feltovich, Michael J Prietula and K Anders Ericsson, 'Studies of Expertise from Psychological Perspectives,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 41; G Deffner, *Think Aloud – An Investigation of the Validity of a Data Collection Procedure* (Peter Lang, 1984); D P Simon and H A Simon, 'Individual Differences in Solving Physics Problems,' in R Siegler (ed) *Children's Thinking: What Develops?* (Erlbaum, 1978) 325; E S Johnson, 'An Information Processing Model of One Kind of Problem Solving' (1964) 4 *Psychological Monographs* 78.
[614] Deffner used 4-second intervals and counted individual letters rather than whole words. See Deffner, above n 613.

Nevertheless, there were a number of incidental findings that appeared to confirm the results recorded by previous researchers. These include, in the context of expertise studies in general:

- Higher-level participants (experts and masters) generally used more effective and efficient risk assessment strategies, with masters being able to correctly evaluate risk based on substantive analysis[615] and recall 'instant' solutions (that is, solutions that appeared to avoid intermediate steps) on this basis.[616]

- There was evidence that higher-level legal specialists were less flexible in their methods[617] and more dependent on contextual information,[618] which may explain the refusals by two expert/master-level participants to provide any assessment because of insufficient background information.

- The atypical case in this study (Case C which involved a completed merger transaction) revealed differences between higher-level participants,[619] who either provided an instant solution based on their expert knowledge or successfully undertook a substantive and conceptually difficult assessment, and on the other hand lower-level participants who either took analytical short cuts (and provided a superficial analysis) or struggled to complete their assessments.

- The responses of some higher-level participants suggested that they had glossed-over unimportant details,[620] while their more effective self-monitoring

[615] K J Gilhooly, P McGeorge, J Hunter, J M Rawles, I K Kirby, C Green and V Wynn, 'Biomedical Knowledge in Diagnostic Thinking: The Case of Electrocardiagram (ECG) Interpretation' (1997) 9 *European Journal of Cognitive Psychology* 199.
[616] W Schneider, 'Training High Performance Skills: Fallacies and Guidelines' (1985) 27(3) *Human Factors* 285.
[617] R J Sternberg and P A Frensch, 'On Being an Expert: A Cost Benefit Analysis,' in R R Hoffman (ed), *The Psychology of Expertise: Cognitive Research and Empirical AI* (Springer Verlag, 1992) 191.
[618] P J Feltovich and H S Barrows, 'Issues of Generality in Medical Problem Solving,' in HG Schmidt and ML de Volder (eds) *Tutorials in Problem-Based Learning* (Van Gorcum, 1984) 128.
[619] As predicted by Norman et al who found in their studies of dermatologists that distinctions between higher-level participants were only evident in their responses to atypical cases. See Geoffrey R Norman, Donald Rosenthal, Lee R Brooks, Scott W Allen and Linda J Muzzin, 'The Development of Expertise in Dermatology' (1989) 125 *Archives of Dermatol*ogy 1063.
[620] J F Voss, G Vesonder and H Spilich, 'Test Generation and Recall by High-Knowledge and Low-Knowledge Individuals' (1980) 19 *Journal of Verbal Learning and Verbal Behavior* 651.

skills[621] could explain why they verbalised fewer comprehension and analytical errors than lower-level participants.[622]

With respect to incidental findings relating specifically to legal expertise, the results reported in this thesis appeared to be confirmatory of previous research in the following areas:

- Higher-level (more expert) participants were faster at diagnosing the legal issues, at least in the sense of more quickly identifying the key factual information in the test-case documents.[623] The area of law itself was known at the start of the tests, so there was no testing of whether or not participants could identify the applicable legal rules.

- The better performance of higher-level participants was likely to have been explained by their possessing more effective mental schemas for assessing legal risk in the context of the test-cases used in this study.[624]

- Higher-level participants had a clearer and more specific idea of what further information they required – and they had more effective strategies for how to get that information. Lower-level experts tended to be more vague in these areas.[625]

- Lower-level participants generally exhibited less precise and more generalised use and recall of information. This was particularly evident in instances of laboured reasoning and in lower-level participants' greater focus on retrieving information which was generally of a lower quality than that relied upon by higher-level participants.[626]

[621] P J Hinds, 'The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Prediction of Novice Performance' (1999) 5 *Journal of Experimental Psychology: Applied* 205.
[622] With respect to the avoidance of comprehension errors, this thesis does not provide reliable evidence on this issue as this was a measure used to rank participants, and hence these errors were assumed to be (but were not proved to be) more likely amongst less expert legal specialists.
[623] Colon-Navarro, above n 604.
[624] Ibid.
[625] Ibid; Weinstein, above n 605.
[626] Weinstein, above n 605.

A critical point of divergence between this thesis and previous studies concerns the underlying explanation for these observed differences. Other researchers have concluded that lower-level participants (typically legal novices in the form of current or recently graduated law students) think differently from higher-level participants because they do not know the area of law as well. This was not the case in the present study where no participant had less than five years specialist experience or required any guidance concerning relevant laws or procedures.

The cognitive differences noted in this thesis were, by design, attributable to factors other than differences in levels of technical legal knowledge. This is reflected in this study's finding that cognitive abilities increase over time even when legal knowledge remains constant, which is an observation that had not been previously demonstrated using the methodologies and experimental techniques used here.

3 *Reassessment*

This thesis indicates that differences in how legal specialists think are evident at each level of development from apprentice to master. This is an important finding inasmuch as these differences have been largely ignored or overlooked in previous research into legal thinking skills. This suggests a need to reassess some past studies, although this issue is likely to be less of a problem in straight novice-expert comparisons and when analysing generic thinking skills such as those relating to reading cases.

It is relatively straightforward to find an individual with a novice-level understanding of an area of law. Accurately identifying and securing the co-operation of expert legal specialists is more challenging. One reason for this is that amongst these latter individuals there can be a significant variation in levels of expertise. Much of the present study was concerned with separating participants into their appropriate expertise category within the same specialist area of law. This has not been the practice in previous research.

Considering first the research of Colon-Navarro,[627] Weinstein[628] and Chay,[629] it is not clear whether or which of their legal specialists were apprentices, journeymen, experts

[627] Colon-Navarro, above n 604.
[628] Weinstein, above n 605.

or masters. By assumption, Colon-Navarro treated his four experts, who had specialised in immigration law for between three and 14 years, to think similarly to each other, but differently from his experienced and inexperienced novices. Weinstein, on the other hand, did not provide details for his three putative experts other than to characterise them as 'outstanding' legal professionals. Chay grouped his two experienced practitioners in the same general category, even though one of them had specialised in family law for only six years and the other had 22 years of family law experience yet did not practice in this area exclusively.

Those researchers who have purported to study how legal experts think when engaging in generic tasks or who have tested for specialist skills but without using think-aloud problem solving or protocol analysis techniques, have also not addressed the assumptions underlying their choices and groupings of experts. For Marchant et al,[630] a tax expert was a tax professional with one to eight years' experience with a multinational accounting firm. For Nievelstein et al,[631] a specialist in civil law was an academic with an average of 5.9 years' post-doctoral experience. For Lundeberg,[632] a legal expert was a lawyer with at least two years' practising or teaching experience.

Such assumptions are arguably less of an issue insofar as many generic skills are likely to have a monotonic correlation to domain expertise such that, in some circumstances, a successful law student may be as good a reader of court cases as a senior partner in a law firm. Nevertheless, a reassessment of these studies may be instructive given the findings of this thesis.

The observation that the cognitive skills of participants in this study appeared to improve according to their levels of expertise while their legal knowledge, in a strictly technical sense, remained constant, may call into question previous research that has purported to demonstrate differences in thinking skills based on variations in legal knowledge amongst participants (the common novice vs expert comparison). In those studies, the observed differences were likely to have been the result of differences in both legal knowledge and cognitive ability. Insofar as this fact was not made clear or

---

[629] Chay, above n 606.
[630] Garry Marchant, John Robinson, Urton Anderson and Michael Schadewald, 'Analogical Transfer and Expertise in Legal Reasoning' (1991) 48 *Organizational Behaviour and Human Decision Processes* 272.
[631] Fleurie Nievelstein, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Expertise-Related Differences in Conceptual and Ontological Knowledge in the Legal Domain' (2008) 20(6) *European Journal of Cognitive Psychology* 1043.
[632] Lundeberg, above n 602.

was overlooked, this thesis provides a basis for reassessing – or at least clarifying – the conclusions of those previous studies.

In the case of Weinstein's subexpert who was a very experienced lawyer but who lacked knowledge of the relevant legal subdomain being tested, his inferior performance compared to experts in the domain was the result of not knowing the law rather than not being able to think like an expert lawyer. Similarly, Mitchell's faculty colleagues were experts in their own fields of law, but they failed to perform as well as their criminal law counterparts when assessing a hypothetical criminal law case.[633] Colon-Navarro's study and Chay's doctoral dissertation, however, relied on comparisons in which both relevant specialist legal knowledge and cognitive ability based on legal experience and expertise were tested. In these studies it may be necessary to reconsider which of these two factors contributed more to the observed differences in performance.

Based on the results of the present study – and in particular the confirmatory findings noted in the previous section – differences in expertise or experience-related cognitive abilities regardless of legal knowledge ostensibly had explanatory significance in relation to Weinstein and Colon-Navarro's findings, and potentially Chay's as well. Whether this points to the overriding effects of generic legal thinking skills or age-related wisdom, or both, is something that future researchers could consider with reference to the discussion presented in this study.

B *Study Limitations*

The aim of this thesis was to find readily identifiable and measurable differences between the cognitive abilities and risk-assessment approaches of legal specialists with different levels of expertise. As an empirical study it required a range of practical as well as theoretical considerations. Some of these considerations led to decisions that limited – in some instances unavoidably – the scope and findings of the study. This is notwithstanding that the key objectives of the thesis were achieved.

The use of protocol analysis to assign participants to groups A, B or C restricted subsequent analysis of their verbal protocols. Having identified instances of exceptional reasoning, conceptual depth and comprehension errors to determine

[633] Mitchell, above n 602.

whether or not some participants would be placed in a particular group, these behaviours could not be later referenced as evidence of performance differences between these groups – at least not directly. In effect, the analysis performed at the ranking stage was an analysis of differences based on previous research on the cognitive skills and abilities associated with different levels of domain expertise. The differences subsequently identified were new and allowed for more nuanced assessments, but were also complementary to, rather than substitutable for, the earlier recorded differences.

This was considered an acceptable trade-off since the practical aims of the study warranted a focus on broader macrocognitive processes and statistically significant differences rather than on piece-meal – and potentially individualistic – verbal protocols.

The methodological approach used to rank participants in this study was described in detail. It was also shown to be dependent on the researcher possessing an in-depth familiarity with this area of law. Without an understanding of the fundamental elements of competition law analysis and an ability to analyse the test-cases beforehand, it would not have been possible to determine workable indicia and weightings for the different categories of analytical approaches that were observed. This contrasts with the measures developed later in the study which constitute a more accessible alternative for determining different levels of specialist legal expertise.

A related issue concerns the exploratory and map-like aspirations of this study. Because the analytical perspective of this thesis was that of a foundational study to be used as a tool for future researchers and for more practical applications, generalisations and the use of aggregated data were both necessary and desirable. Having sought to establish an essentially generic basis for distinguishing different levels of expertise using the traditional categories of apprentice, journeyman, expert and master, some more subtle differences were inevitably overlooked.

For instance, where just one or two participants in group A, B or C demonstrated a particular cognitive skill, this would likely not have met the threshold requirement of statistical significance when comparing different groups. Any number of such differences could have been discussed. But their utility in the broader context of the thesis, which sought to use broad strokes to highlight different legal-risk assessment

behaviours with potential application in other specialist areas of law, was viewed as limited.

This is not to say that more fine-grained analyses could not have been undertaken. Rather, this was not the approach chosen given the research question that had been posed and the limits on the length of this thesis. As noted elsewhere in this chapter, more detailed analyses have not been precluded in terms of the data recorded in this study or further research projects. Indeed, seen in the context of the overall aims of this thesis, such further research can be expected to benefit from the results and findings presented here, particularly insofar as sample-groups of legal specialists may now be more confidently ranked according to their levels of expertise.

The initial ranking of participants in this study relied on a form of differential analysis in which only those participants who could be distinguished from the main group of participants could be confidently categorised as experts and masters (Group A) or apprentices and journeymen (Group C). This approach relied on a number of assumptions about the sample of volunteers who participated in the study and the appropriateness of a presumed normal distribution of levels of expertise. Apparent concerns were to a degree ameliorated by only ascribing *relative levels of likely expertise* to individual participants, rather than more conclusive scores of absolute expertise. Moreover, within each group no further intra-group ranking was attempted. Yet there remains the possibility that some participants were categorised either higher than their actual levels of expertise or, the more likely scenario given the conservative nature of the ranking process, lower than their actual levels of expertise.

A further price paid under this approach was the restricted usability of Group B participants. Because these participants were not separable according to their levels of likely expertise – which was unavoidable given the comparative, clustering approach adopted for initial ranking purposes – observed differences between these participants both internally and with the other two groups would be necessarily inconclusive. The main value of the legal-risk assessments performed by these participants was as descriptive examples of different reasoning behaviours. This was on one view a design inefficiency and one which, even if ultimately unavoidable, resulted in the suboptimal utilisation of study volunteers.

A related issue was the size of the Group B category, and specifically whether it was too broadly defined. On the one hand, it was important that Group A and Group C participants be separated by more than the results of a single measure of likely expertise. This required at least five-points of separation. On the other hand, this degree of separation meant that Group B contained eight participants who undertook a total of 30 legal-risk assessments. Whether this gap could have been narrowed with a more refined scoring and weighting system is arguable. Within the limitations of the system as conceived it was ostensibly as narrow as the underlying assumptions would permit. Another study along similar lines may be able to revisit and refine both the relevant assumptions and related scoring and weighting methodologies.

Another concern was that the breadth of the Group B category resulted in it containing most of the expert and journeyman participants while leaving Group A to consist mostly of masters and Group C mostly apprentices. Subsequent analysis would therefore have been between more distantly related legal specialists which could introduce material comparative discontinuities. This could explain why in Chart 6.2 there was a substantial gap between apprentices and journeymen on the one hand and experts and masters on the other in terms of time spent engaging in synthesis. However, Chart 6.1 supports the counter-argument that there was a clear progressive difference in terms of time spent retrieving information from long-term memory and that on this measure journeymen and experts were indistinguishable, which suggested a contiguous alignment.

Considerable thought was given to the use of a second assessor to confirm the categorisation of particular behaviours and the identification of certain cognitive tasks. There was clear precedent for using such confirmation, but this has usually been because it is demanded by the rigour of a particular kind of scientifically-grounded analysis or was simply a researcher's default position. The choice of not using a second assessor for this thesis was supported by both practical and theoretical precedent, which seemed more applicable to the present study for the reasons stated in Chapter 5. The trade-off was a need to provide more detailed descriptions of the assessment process rather than rely on a simple statement of inter-rater agreement. Moving forward from the present map-like, exploratory approach of cognitive analysis to more focused research projects investigating a smaller number of more discretely-

defined variables, the balance of considerations is likely to fall towards the use multiple assessors to provide the assurances expected of such projects.

The inherent inefficiency of the chosen investigatory approach also needs mentioning. While the results and findings described in this thesis have been presented systematically, the process by which they were selected was iterative and at times circuitous. There were numerous attempts over many months to investigate different aspects of participants' transcripts using a variety of qualitative and quantitative analytical techniques. Some of these looked promising at first, but the end results were either not compelling or overly-complicated. Others were dependent on too many assumptions. Had the presented approach been selected initially, many of these dead-end investigations and prolonged iterations would have been avoided. Fortunately, it was the researcher alone who experienced this excessive exertion (since all interviews were completed and transcribed at an early stage study participants were spared this extra work) and, in any event, these processes were ultimately viewed as crucial, if not defining, aspects of the chosen approach.

Lastly, it may have been more useful to have asked participants to assign a percentage estimate to the likelihood that a particular merger transaction would be cleared by the relevant competition authority. While it was still possible to identify how those participants who provided conclusive opinions viewed the likely outcome of cases, this involved a binary distinction rather than a graduated scale of degree. Moreover, those participants who gave inconclusive assessments may have been able to give a percentage figure, albeit a heavily qualified one. This could have facilitated a more detailed analysis of assessment accuracy. Nevertheless, it would still have been necessary to understand the reasoning behind any such percentage estimates to determine the extent to which they were logically-based and compelling, rather than simply the result of guesswork.

The main downside to requiring a percentage estimate of the likely outcome of a case is that it could have obscured or masked an otherwise inconclusive assessment. If all participants strove to provide a conclusive assessment in terms of a percentage estimate, it would likely have been more difficult to determine which of their assessments were in fact inconclusive. This could have been a significant hindrance to

identifying inconclusive assessments, which were found to be more prevalent amongst apprentice and journeyman-level participants than higher-level participants.

C *Future Research*

Using methodologies and previous research from the field of cognitive psychology, this thesis outlined a number of qualitative and quantitative methods for determining the expertise levels of legal specialists. As the first study of its kind, it requires confirmation through replication of its results and findings. If that occurs, then the conclusions drawn in the previous chapter may become foundational to further research in a range of other areas, including: the cognitive development of legal expertise; the formal accreditation of legal specialists; assessments by consumers of legal services; the development of legal talent within law firms; and legal education, particularly at intermediate and advanced levels.

1 *Replication and Reproduction of Results*

Detailed descriptions of the methodologies used in this thesis were provided in Chapter 3 to enable and encourage the reproduction of this study. Not only were the testing and analytical techniques (and technologies) described and explained at length, the actual test-case materials and website used to guide the test process have been preserved on-line.[634] Given that these particular cases could be quickly forgotten amongst the general population of Australian and New Zealand competition law specialists, it may be feasible to use the same cases to attempt the direct reproduction of the results and findings described here.

Alternatively, or in addition to any such future studies, research could be undertaken along similar lines in other areas of law. In this respect, the qualitative and quantitative data generated in the present study are not specific only to competition law expertise. Moreover, the test techniques, technologies and even the test website used in this study could be easily adapted for investigations in other specialist legal fields.

A summary of the methodological approach behind this thesis and links to the relevant test materials and test website were provided in Part F of Chapter 3.

[634] http://taps28nztest.wordpress.com/ and http://taps28info.wordpress.com/

2 *Cognitive Development of Legal Expertise*

The focus of this thesis was on readily identifiable and measurable differences between how different levels of legal experts think when assessing legal risk in the same information-limited and time-constrained contexts. This resulted in data that were intentionally generalizable and which were derived from analytical approaches chosen for their ready applicability and statistical significance. Given limits on the length of discussion in this thesis, priority was given to more broadly applicable measures over the detailed analysis of individual protocols.[635]

This presents an opportunity for further research at this next level of analysis, which in turn could lead to a better understanding of the cognitive stages or steps through which legal expertise develops as individuals progress from apprentice to journeyman, and then to expert and master. Assuming that further studies are able to confirm the findings of the present study, researchers seeking to study cognitive development within this process can use the results presented here to identify and categorise future participants in such studies.

The discussion in Part D of the previous chapter presented a contextual analysis of how the concepts and theories relating to System 1 and System 2 thinking could provide opportunities for further investigation into why study participants with different levels of expertise thought as they did. This is one area of opportunity for further studies of the cognitive development of legal expertise. The results of that analysis were useful both contextually for this study and as a starting point for different kinds of cognitive analysis.

The contribution of this thesis to such further research could be substantial insofar as it provides a workable basis for ranking different levels of legal specialists. Prior to the present study, researchers relied on a number of critical assumptions, some of which may now be considered questionable. These studies have also most commonly compared novices and assumed experts, which has precluded more fine-grained analyses at the higher-end of the legal expertise spectrum.

[635] Detailed protocol analysis was, however, integral to ranking participants according to their demonstrated abilities to engage issues with conceptual depth, undertake exceptionally good or bad reasoning, and avoid comprehension errors. As noted further above, this largely precluded detailed protocol analysis in subsequent stages of the study to avoid double counting and conflating evidence with assumptions.

As previously noted, the methodology for ranking participants in this study involved detailed protocol analysis. This increased the risk that further detailed protocol analysis when comparing different levels of experts would result in double-counting the same cognitive attributes. For this reason, a broader and less detailed (but ultimately more purposeful) approach was adopted for this study's back-end analysis. Future researchers can avoid these issues by starting with the behavioural traits and cognitive indicators identified in this thesis to rank their participants, after which they will have greater freedom to examine and compare their detailed protocols.

3 *Formal Accreditation of Legal Specialists*

Chay selected the higher-level participants in his study based on their having passed the Queensland Law Society's family-law specialist accreditation process.[636] Like many specialist legal accreditation schemes around the world, the Queensland Law Society requires that in addition to successfully completing set tasks, such as mock client-interviews, prospective applicants must have a minimum number of years' experience in the relevant area of law (at least 5 years in the case of Chay's experts) and be 'recognised by her or his peers as a competent family law practitioner.'[637] The Law Society of England and Wales states with respect to its specialist accreditation scheme, 'The expertise of scheme members in a given area of law has been verified and their accomplishment is recognised by the Law Society.'[638] Similarly, the State Bar of California states that its certified specialists:

> Have taken and passed a written examination in their specialty field, demonstrated a high level of experience in the speciality field [a minimum 25% of their time for each the previous 5 years], fulfilled ongoing education requirements and been favourably evaluated by other attorneys familiar with their work.[639]

These accreditation and certification schemes do not utilise a cognitive-based methodology for testing applicants, nor do they rank them according to their apparent levels of specialist legal expertise. They also typically focus only on broad legal specialities rather than subspecialties, and primarily only on consumer-related practice

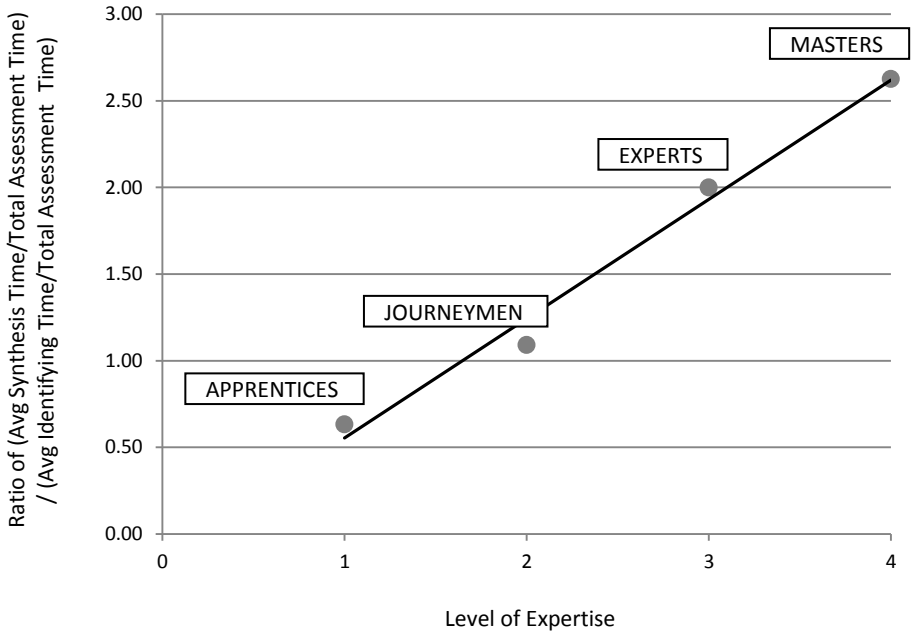
---

[636] Chay, above n 606, 82.
[637] Ibid 81.
[638] The Law Society, Accreditation http://www.lawsociety.org.uk/support-services/accreditation/ accessed 1 December 2014.
[639] The State Bar of California, Legal Specialization, Frequently Asked Questions http://ls.calbar.ca.gov/LegalSpecialization/FAQ.aspx accessed 1 December 2014.

areas rather than corporate law. Of interest in this particular context are the data from the previous chapter which can be used to assist with the ranking of legal specialists under existing accreditation schemes.

Chart 7.1 below makes use of the quantitative data reflected in Column 3 of Table 6.1 in the previous chapter. That column, which lists cognitive indicators for each of the above levels of expertise, reveals different ratios of synthesis time and identifying time for each of these expert levels. While based on sample averages, the linear trend line in this chart has explanatory power and is statistically significant.[640]

CHART 7.1 – Ratio of Average Synthesis Time to Average Identifying Time[641]



This chart combines the information presented in charts 6.1 and 6.2 in Chapter 6 (as summarised in Column 3 of Table 6.1). The result is a ranking of expertise based solely on time spent synthesising and identifying issues when assessing legal risk in an information-limited and time-constrained context. With further work, this chart could be augmented by the additional information identified in this thesis regarding different types of assessment behaviour (ease of reasoning, certainty of assessment, assessment accuracy and depth of analysis as discussed in Chapter 5) and with reference to verbalisation rates while engaging in different cognitive tasks.

[640] This regression line has an Adjusted $R^2 = 0.979$ and a p-value = 0.007.
[641] The actual ratio – ie the ratio for which statistically significant data are available – is as indicated on the vertical axis of this chart = (Average Synthesis Time/Total Assessment Time)/(Average Identifying Time/Total Assessment Time). The above simplified title is for presentation purposes only.

Future research could assess the extent to which these data ratios and additional behavioural information could be used as a basis for – or at least assist in – formally accrediting legal specialists according to their levels of sub-domain expertise.

4 *Assessments by Consumers of Legal Services*

According to a 2013 survey of 968 in-house legal counsel, the most common reason given for selecting a particular law firm or legal advisor was their expertise in a specific field of law.[642] This finding is consistent with the previously mentioned schemes of specialist accreditation, by which lawyers seek to gain 'brand advantage through establishing themselves as a specialist expert in a given area of law.'[643] Apart from such accreditation (which is of limited relevance for larger corporate clients), users of legal services have little to rely on when assessing legal specialists and nothing that provides a cognitive-based ranking of individual lawyers' according to their levels of specialist legal expertise.

Whether or not a standardised testing methodology using cognitive data is developed by law societies or other formal accreditation bodies, some larger users of legal services could develop their own proprietary methodology to rank individual lawyers using information like that shown in Table 6.1 and Chart 7.1. Such a methodology could also incorporate accreditation information and data provided by publications such as *Who's Who Legal*, *Chambers and Partners* and *The Legal 500* as discussed in Chapter 4. However, as previously noted these publications do not provide quantitative rankings as such, and nor do they provide information on apprentice and journeyman-level legal specialists.

These are issues that future researchers could consider, potentially with the support of legal consumers and interested law firms.

---

[642] Acritas, *Winning and Losing Business: Clients' Candid Views on Why They Hire and Fire Their Law Firms* (Acritas, September 2013). This survey found that 37% of interviewees identified subject-matter expertise as the main reason they hired particular external legal advisors. The other reasons were the geographic presence of the advisor (17%), a previous positive experience with the advisor or their firm (15%), and the advisor's cost effectiveness (11%).

[643] The Law Society of England and Wales, Specialist Schemes – Member Benefits, http://www.lawsociety.org.uk/support-services/accreditation/specialist-schemes/ accessed on 1 December 2014.

5 *Development of Legal Talent Within Law Firms*

Mentoring, coaching and improvements in feedback for mid-level lawyers are all subjects that exercise the minds of law firm management.[644] While the motivations behind initiatives in these areas can be complex, a central challenge is how to assess and develop the talent of legal staff, particularly those at the mid to senior-levels who have already mastered the rules of legal reasoning and generic skills taught in law schools. By engaging them in a discussion about the cognitive skills and behaviours discussed in this thesis, it may be possible to identify new strategies for increasing individual lawyer's expertise in novel and effective ways.

As an input into a cognitive apprenticeship model, an understanding of the results and findings of this study could assist focus and guide training programs as well as inform individual development plans for legal specialists at various levels. During the interviews conducted for this study, there was considerable interest in this area amongst both partner and non-partner level participants. Indeed, without a degree of fascination with and belief in the usefulness of this type of research, it is doubtful that the number of participants who volunteered for this study would have been as large as it was.

Further research into how to develop the findings of this thesis to better equip those with senior management responsibility within law firms would ostensibly be a welcome and worthwhile pursuit.

6 *Legal Education*

Law schools have traditionally focused on teaching students to think like lawyers, but not on preparing legal specialists to think like experts. Using the methodologies and findings from this thesis, law schools can review current post-graduate programs which focus on a specialist field of law and target mid to senior-level lawyers.[645] This thesis has shown that levels of expertise vary materially even amongst experienced legal specialists, and that this variation is evident in legal risk-assessment behaviours and

[644] See, for instance, Nickolas John James, 'Professional Mentoring Programs for Law Students' (2011) 30 *University of Tasmania Law Review* 90; Leslie A Gordon, *Legal Wranglers: Law Firms Add Coaches to their Staffs*, ABA Journal, May 2014; William D Henderson, *Supercharging Lawyer Development Through Feedback*, NALP Bulletin, June 2014.

[645] An example of such programs are the Specialist LLM (Master of Laws) programs offered by over 80% of the 86 world's top providers of LLM courses as ranked by the Financial Times in its October 2012 survey at http://rankings.ft.com/lawschools/llm-2012-listing accessed 1 December 2014.

cognitive performance on tasks such as identifying issues and engaging in synthesis in information-limited and time-constrained contexts.

The testing methods used in this study were demonstrated to be effective in eliciting verbal data relevant to how participants approached the assessment of issues in the test cases. Moreover, these tests could not have been gamed or pre-empted by participants in the way that some forms of assessment at law schools can be undermined by students who, for instance, engage third-parties to complete their assignments or cheat the system in other ways.[646] Even if an interviewee knows that reduced identifying time and increased synthesis time generally reflects a higher level of expertise, without the ability to actually identify relevant issues quickly and synthesise them effectively, this information is of little help. In addition, this method of testing enables the pinpointing of specific areas of cognitive difficulty that may not be evident using traditional examination techniques.

In terms of essay writing, Palasota has shown how think-aloud problem solving can be adapted to understand the thinking processes of students, although verbalising while writing an essay involves different considerations from those noted in this thesis.[647] Colon-Navarro[648] and Chay[649] have also demonstrated that verbal data can be elicited during the review of video-recorded interview sessions. Further uses of think-aloud verbalisations in law school settings have been suggested by Senger.[650]

These represent other areas of research that could be undertaken by both legal scholars and researchers outside the legal academy such as other educators and cognitive psychologists. Insofar as such research focuses on higher-level legal specialists, this thesis offers new methodological guidance and previously unavailable comparative data.

---

[646] A recent example of a third-party system that has raised concerns of these kinds is MyMaster, which has been described as 'an underground essay writing business … [that] reaped more than $160,000 from students studying in Australia [in 2014].' Amy McNeilage, 'Universities Investigate MyMaster Cheating,' *Sydney Morning Herald*, 18 November 2014 http://www.smh.com.au/national/education/universities-investigate-mymaster-cheating-20141118-11p33f.html.

[647] Anthony Palasota, 'Expertise and the Law: Some Recent Findings from the Cognitive Sciences About Complex Human Information Processing' (1990-1991) 16 *Thurgood Marshall Law Review* 599.

[648] Colon-Navarro, above n 604.

[649] Chay, above n 606.

[650] Charles J Senger, 'Thinking Aloud Protocols: A Diagnostic Tool for Teaching Legal Problem Solving' (1993) 10 *Thomas M Cooley Law Review* 367 (see in particular, Part IV: Law School Applications, 378-381).

D *Conclusion*

Each of the 20 volunteers who participated in this study had at least five years' experience as a competition law specialist. Together they generated over 70,000 words of think-aloud verbalisations. These verbalisations reflected the cognitive processes of participants as they engaged in the assessment of legal risk in merger clearance cases that were at the time being considered by either the ACCC or CC. The selection of these volunteers and these cases (and the rules for disqualifying cases when a participant had direct prior knowledge of a test-case transaction) ensured that no participant had a knowledge advantage in terms of the relevant laws and procedures, of the factual details of particular transactions, or of the outcome of the ACCC or CC's deliberations, which were the central focus of the legal-risk assessment task.

Through the application of previous research findings concerning the cognitive abilities of experts, participants were ranked and their assessments divided into the categories of apprentice, journeyman, expert and master-level according to Hoffman's Scheme. This progressive domain-specific scheme of expertise development was chosen over the knowledge-based categorisations of expertise used by researchers such as Patel and Groen, whose analytical framework has been the predominant influence in the design of previous empirical studies of legal thinking skills.

The exploratory nature of the analysis of verbal protocols undertaken in this study was a further point of departure from previous research in this area in a methodological sense, although there was firm precedent for this approach in the work of Baddely and other scholars. The results of this exploration were both qualitative and quantitative, enabling for the first time a listing of readily identifiable and measurable differences between different levels of legal experts – differences that were unrelated to variations in technical legal knowledge. This record of observed differences opens up opportunities for further research in a number of different areas, as well as the possibility to confirm the findings presented here through the recreation of the test procedures and methodological choices described in detail in the preceding chapters.

Ultimately, it is hoped that these contributions to existing scholarship on how lawyers think will enhance our effectiveness as lawyers and law teachers, and lead to opportunities for more targeted research into the cognitive skills of legal experts both within and beyond the legal academy.

BIBLIOGRAPHY

A *Articles/Books/Reports*

Aaronson, Mark Neal, 'We Ask You to Consider: Learning About Practical Judgment in Lawyering' (1997-1998) 4 *Clinical Law Review* 251

Ackerman, Phillip L, 'Nonsense, Common Sense and Science of Expert Performance: Talent and Individual Differences' (2013) 45 *Intelligence* 6

Acritas, *Winning and Losing Business: Clients' Candid Views on Why They Hire and Fire Their Law Firms* (Acritas, September 2013)

Alexander, Larry, *Demystifying Legal Reasoning* (Cambridge University Press, 2008)

Alexander, P A, 'Can We Get There From Here?' (2003) 32 *Educational Researcher* 3

Amsel, Eric, Rosanna Langer and Lynn Loutzehiser, 'Do Lawyers Reason Differently From Psychologists? A Comparative Design for Studying Expertise' in R Sternberg and P Frensch (eds), *Complex Problem Solving: Principles and Mechanisms* (Lawrence Erlbaum Associates, 1991)

Amsterdam, Anthony G, 'Clinical Legal Education – A 21st Century Perspective' (1984) 34 *Journal of Legal Education* 612

Ardelt, Monika, 'Wisdom as Expert Knowledge Systerm: A Critical Review of a Contemporary Operationalization of an Ancient Concept' (2004) 47 *Human Development* 257

Arocha, Jose F and Vimla L Patel, 'Novice Diagnostic Reasoning in Medicine: Accounting for Evidence' (1995) 4(4) *The Journal of the Learning Sciences* 355

Australian Competition and Consumer Commission, *Informal Merger Review Process Guidelines* (Australian Competition and Consumer Commission, September 2013)

Australian Competition and Consumer Commission, *Brambles Industries Limited – Completed Acquisition of CEVA Limited's Pallecon Business* (Reference 50545), Market Inquiries Letter, 19 March 2013

Australian Competition and Consumer Commission, *Ruralco Holdings Limited – Proposed Acquisition of Elders Rural Services Limited* (Reference 50978), Market Inquiries Letter, 19 March 2013

Australian Competition and Consumer Commission, *Toll Holdings Limited – Proposed Acquisition of Linfox Trans-Bass Business* (Reference 50982), Market Inquiries Letter, 22 March 2013

Australian Competition and Consumer Commission, *Rocla Pty Ltd – Proposed Acquisition of Beresford Concrete Products Pty Ltd* (Reference 51157), Market Inquiries Letter, 23 April 2013

Australian Competition and Consumer Commission, *ACCC To Not Oppose Ruralco's Proposed Acquisition of Elders Rural Services*, Media Release, 30 May 2013

Bachmann, T and M Oit, 'Stroop-like Interference in Chess Players' Imagery: An Unexplored Possibility to be Revealed by the Adapted Moving-Spot Task' (1992) 54 *Psychological Research* 27

Baddeley, A, 'Short-Term and Working Memory,' in E Tulving and F Craik (eds), *The Oxford Handbook of Memory* (Oxford University Press, 2000) 77

Baddely, A, 'Is Working Memory Still Working?' (2002) 7(2) *European Psychologist* 85

Baddeley, A, 'Working Memory: Theories, Models, and Controversies' (2012) 63 *Annual Review of Psychology* 1

Baltes, P B and U M Staudinger, 'Wisdom: A Metaheuristic (Pragmatic) to Orchestrate Mind and Virtue Toward Excellence' (2000) 55 *American Psychologist* 122

Baltes, P B, U M Staudinger, A Maercker and J Smith, 'People Nominated as Wise: A Comparative Study of Wisdom-Related Knowledge' (1995) 10(2) *Psychology and Aging* 155

Baxt, R, 'The Australian Concept of Market – How It Came To Be,' in M Richardson and P Williams (eds), *The Law and The Market* (The Federation Press, 1995) 10

Beaton-Wells, Caron, *Proof of Antitrust Markets in Australia* (The Federation Press, 2003)

Bedard, J, 'Expertise and Its Relation to Audit Decision Quality' (1991) 8 *Contemporary Accounting Research* 198

Bilalic, Merim, Peter McLeod and Fernand Gobet, 'Specialization Effect and Its Influence on Memory and Problem Solving in Expert Chess Players' (2009) 33 *Cognitive Science* 1117

Blasi, Gary L, 'What Lawyers Know: Lawyering Expertise, Cognitive Science, and the Functions of Theory' (1995) 45(3) *Journal of Legal Education* 313

Blasi, Gary L, 'Teaching/Lawyering as an Intellectual Project' (1996) 14 *Journal of Professional Legal Education* 65

Bonner, Sarah M and Jerome V D'Agostino, 'A Substantive Analysis of Responses to Items from the Multistate Bar Examination' (2012) 25 *Applied Measurement in Education* 1

Bonneville-Roussy, A, G L Lavigne and R J Vallerand, 'When Passion Leads to Excellence: The Case of Musicians' (2011) 39 *Psychology of Music* 123

Borst, Gregoire, Elaine Niven and Robert H Logie, 'Visual Mental Image Generation Does Not Overlap With Visual Short-Term Memory: A Dual-Task Interference Study' (2012) 40 *Memory & Cognition* 360

Boshuizen, Henny P A and Henk G Schmidt, 'On the Role of Biomedical Knowledge in Clinical Reasoning by Experts, Intermediates and Novices' (1992) 16(2) *Cognitive Science* 153

Braten, Ivar and Helge I Stromso, 'A Longitudinal Think-Aloud Study of Spontaneous Strategic Processing During the Reading of Multiple Expository Texts' (2003) 16 *Reading and Writing: An Interdisciplinary Journal* 195

Brooks, David, *The Social Animal: The Hidden Sources of Love, Character and Achievement* (Random House, 2012)

Brown, A and J Day, 'Macro Rules for Summarizing Strategies: The Development of Expertise' (1983) 22 *Journal of Verbal Learning and Verbal Behaviour* 1

Brozek, Bartosz and Jerzy Stelmach, *Methods of Legal Reasoning* (Law and Philosophy Library, 2006)

Brunt, M, '"Market Definition" Issues in Australian and New Zealand Trade Practices Litigation' (1990) 18 *Australian Business Law Review* 86

Bryan, W L and N Harter, 'Studies on the Telegraphic Language: The Acquisition of a Hierarchy of Habits' (1899) 6 *Psychological Review* 345

Bryden, David P, 'What Do Law Students Learn? A Pilot Study' (1984) 34 *Journal of Legal Education* 479

Butterworth, J Scott and Edmund H Reppert, 'Auscultatory Acumen in the General Medical Population' (1960) 174(1) *JAMA* 114

Cacciabue, P C and E Hollnagel, 'Simulation of Cognition: Applications,' in J M Hoc, P C Cacciabue and E Hollnagel, *Expertise and Technology: Cognition and Human-Computer Cooperation* (Lawrence Erlbaum Associates, 1995) 55

Camp, Elisabeth, 'Metaphor in the Mind: The Cognition of Metaphor' (2006) 1(2) *Philosophy Compass* 154

Chadderton, Sam, 'Time for Competition Law to "Get Real," Says Incoming CAT Chairman,' *The Lawyer*, 11 January 2013

Chaiken, S and Y Trope (eds), *Dual-Process Theories in Social Psychology* (Guilford Press, 1999)

Charles River Associates, *Benchmarking the Supply of Legal Services by City Law Firms*, report prepared for the UK Legal Services Board (Charles River Associates, 2011)

Charness, N, 'The Impact of Chess Research on Cognitive Science' (1992) 54 *Psychological Research* 4

Charness, N, M Tuffiash and T Jastrzembski, 'Motivation, Emotion, and Expert Skill Acquisition,' in D Dai and RJ Sternberg (eds), *Motivation, Emotion and Cognition: Integrative Perspectives* (Erlbaum, 2004) 299

Chase, W G and H A Simon, 'Perception in Chess' (1973) 4 *Cognitive Psychology* 55

Chase, W G and H A Simon, 'The Mind's Eye in Chess,' in WG Chase (ed), *Visual Information Processing* (Academic Press, 1973) 215

Chay, Allan James, *Lawyer Problem Solving: An Investigation of the Knowledge Used in Solving Practical Legal Problems* (PhD Thesis, Griffith University, 2006)

Chi, M T H, 'Knowledge Structure and Memory Development,' in R Siegler (ed) *Children's Thinking: What Develops?* (Erlbaum, 1978) 73

Chi, M T H, P Feltovitch and R Glaser, 'Categorization and Representation of Physics Problems by Experts and Novices' (1981) 5 *Cognitive Science* 121

Chi M T H, R Glaser and E Rees, 'Expertise in Problem Solving,' in R Sternberg (ed) *Advances in the Psychology of Human Intelligence* (Vol. 1) (Erlbaum, 1982) 7

Chi, Michelene T H, 'Theoretical Perspectives, Methodological Approaches, and Trends in the Study of Expertise,' in Y Li and G Kaiser (eds) *Expertise in Mathematics Instruction* (Springer, 2011) 17

Chi, Michelene T H, 'Laboratory Methods for Assessing Experts' and Novices' Knowledge,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 167

Chi, M T H, 'Two Approaches to the Study of Experts' Characteristics,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman, *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 21

Christensen, Leah, 'The Psychology Behind Case Briefing: A Powerful Cognitive Schema' (2006) 29(1) *Campbell Law Review* 5

Christensen, Leah M, 'Legal Reading and Success in Law School: An Empirical Study' (2006-2007) 30 *Seattle University Law Review* 603

Christensen, Leah M, 'The Paradox of Legal Expertise: A Study of Experts and Novices Reading the Law' (2008) *Brigham Young University Education and Law Journal* 53

Clarke, Julie, *The International Regulation of Transnational Mergers* (PhD Thesis, Queensland University of Technology, 2010)

Colon-Navarro, Fernando, 'Thinking Like a Lawyer: Expert-Novice Differences in Simulated Client Interviews' (1997) 21 *The Journal of the Legal Profession* 107

Colvin, Geoffrey, *Talent is Overrated: What Really Separates World-Class Performers from Everybody Else* (Portfolio Trade, 2010)

Commerce Commission, *Determination: Bertelsmann SE & Co KGaA and Pearson plc* [2013] NZCC 6 (Wellington, 19 March 2013)

Commerce Commission, *Determination: Bligh Finance Limited and Hire Equipment Group Limited* [2013] NZCC 2 (Wellington, 21 February 2013)

Commerce Commission, *Statement of Preliminary Issues: Hirepool / Hirequip* (Wellington, 22 October 2012)

Commerce Commission, *Statement of Preliminary Issues: Penguin / Random House* (Wellington, 11 January 2013)

Commerce Commission, *Statement of Preliminary Issues: Perry Metal Protection Limited / CSP Coating Systems* (Wellington, 25 January 2013)

Coyle, Daniel, *The Talent Code: Greatness Isn't Born. It's Grown. Here's How* (Bantam, 2009)

Crandall, Beth, Gary Klein and Robert R Hoffman, *Working Minds: A Practitioner's Guide to Cognitive Task Analysis* (The MIT Press, 2006)

Crombag, H F M, J L De Wijkerslooth and E H Van Tuyl Van Serooskerken, 'On Solving Legal Problems' (1975-1976) 27 *Journal of Legal Education* 168

Cutler, Anne, *Encyclopedia of Cognitive Science* (John Wiley & Sons, 2005)

Dancey, C and J Reidy, *Statistics Without Maths for Psychology: Using SPSS for Windows* (Prentice Hall, 2004)

Dane, Erik Ian, *Does Experience Matter? Examining the Relationship Between Experience and Dynamic Decision-Making Effectiveness Among Professionals* (PhD Thesis, University of Illinois at Urbana-Champaign, 2007)

Dane, Erik, Kevin W Rockmann and Michael G Pratt, 'When Should I Trust My Gut? Linking Domain Expertise to Intuitive Decision-Making Effectiveness' (2012) 119 *Organizational Behavior and Human Decision Processes* 187

Dawes, R M, 'A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making' (1971) 26 *American Psychologist* 180

Deegan, Dorothy H, 'Exploring Individual Differences Among Novices Reading in a Specific Domain: The Case of Law' (1995) 30(2) *Reading Research Quarterly* 154

Deffner, G, *Think Aloud: An Investigation of the Validity of a Data Collection Procedure* (PhD Thesis, University of Hamburg, 1983)

Deffner, G, *Think Aloud – An Investigation of the Validity of a Data Collection Procedure* (Peter Lang, 1984)

de Groot, A, *Thought and Choice in Chess* (Mouton, 1965)

Dijkstra, K A, J van der Pligt, G A van Kleef and J H Kerstholt, 'Deliberation Versus Intuition: Global Versus Local Processing in Judgment and Choice' (2012) 48 *Journal of Experimental Social Psychology* 1156

Djkstra, Koen A, Joop van der Pligt and Gerben A van Kleef, 'Deliberation Versus Intuition: Decomposing the Role of Expertise in Judgment and Decision Making' (2013) 26 *Journal of Behavioral Decision Making* 285

Dreyfus, H L and S E Dreyfus, *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer* (The Free Press, 1986)

Duckworth, A L, T A Kirby, E Tsukayama, H Berstein, and K A Ericsson, 'Deliberate Practice Spells Success: Why Grittier Competitors Triumph at the National Spelling Bee' (2012) 2 *Social Psychology and Personality Science* 174

Dudley, Robert J, *Think Like a Lawyer: How to Get What You Want by Using Advocacy Skills* (Nelson-Hall, 1980)

Duncker, K, 'A Qualitative (Experimental and Theoretical) Study of Productive Thinking (Solving Comprehensible Problems)' (1926) 33 *Pedagogical Seminary* 642

Duncker, K, *On Problem Solving* (Greenwood Press, 1972 (L S Lees Trans.)) (Original work published 1945)

Durkin, H E, 'Trial and Error, Gradual Analysis and Sudden Reorganization: An Experimental Study of Problem Solving' (1937) *Archives of Psychology* 210

Eells, Tracy D, Kenneth G Lombart, Edward M Kendjelic, L Carolyn Turner and Cynthia P Lucas, 'The Quality of Psychotherapy Case Formulations: A Comparison of Expert, Experienced, and Novice Cognitive-Behavioral and Psychodynamic Therapists' (2005) 73(4) *Journal of Consulting and Clinical Psychology* 579

Elhauge, E and D Geradin, *Global Antitrust Law & Economics* (Foundation Press, 2007)

Elhauge, Einer, 'How Should Competition Law Be Taught?' (2008) 4(1) *Competition Policy International* 267

Elo, A E, 'Age Changes in Master Chess Performance' (1965) 20 *Journal of Gerontology* 289

Elo, A E, *The Rating of Chessplayers, Past and Present* (Arco Chess, 2nd ed, 1986)

Elstein, Arthur S, 'Naturalistic Decision Making and Clinical Judgment' (2001) 14(5) *Journal of Behavioral Decision Making* 363

Epstein, S, 'Integration of the Cognitive and Psychodynamic Unconcious' (1994) 49 *American Psychologist* 709

Ericsson, K Anders, 'Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains' (2004) 79(10) *Academic Medicine* S70

Ericsson, Anders K and Herbert Simon, *Protocol Analysis: Verbal Reports as Data* (The MIT Press, 1984/93)

Ericsson, K Anders, 'An Introduction to Cambridge Handbook of Expertise and Expert Performance: Its Development, Organization and Content' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 3

Ericsson, K Anders, 'Protocol Analysis and Expert Thought: Concurrent Verbalizations of Thinking during Experts' Performance on Representative Tasks' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 223

Ericsson, K Anders, 'The Influence of Experience and Deliberate Practice on the Development of Superior Performance,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 683

Ericsson, K Anders, 'The Scientific Study of Expert Levels of Performance: General Implications for Optimal Learning and Creativity' (1998) 9(1) *High Ability Studies* 75

Ericsson, K A, R T Krampe and C Tesch-Romer, 'The Role of Deliberate Practice in the Acquisition of Expert Performance' (1993) 100 *Psychological Review* 363

Ericsson, K A and A C Lehman, 'Expert and Exceptional Performance: Evidence on Maximal Adaptations on Task Constraints' (1996) 47 *Annual Review of Psychology* 273

Evans, Vyvyan, 'Cognitive Linguistics' (2012) 3 *WIREs Cognitive Science* 129

Farrar, John, *Legal Reasoning* (Thomson Reuters, 2010)

Fels, Allan, Sharon Henrick and Martyn Taylor, 'Australia,' in Maher Dabbah and Paul Lasok (eds) *Merger Control Worldwide* (Cambridge University Press, 2012) 78

Feltovich, P J, P E Johnson, J M Moller and D B Swanson, 'LCS: The Role and Development of Medical Knowledge in Diagnostic Expertise,' in William J Clancey and Edward H Shortliff (eds), *Readings in Medical Artificial Intelligence: The First Decade* (Addison-Wesley, 1984)

Feltovich, P J and H S Barrows, 'Issues of Generality in Medical Problem Solving,' in HG Schmidt and ML de Volder (eds) *Tutorials in Problem-Based Learning* (Van Gorcum, 1984) 128

Feltovich, Paul J, Michael J Prietula and K Anders Ericsson, 'Studies of Expertise from Psychological Perspectives,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 41

Fernandes, Ronald and Herbert A Simon, 'A Study of How Individuals Solve Complex and Ill-Structured Problems' (1999) 32 *Policy Sciences* 225

Fortescue, Michael, *Neural Network Model of Lexical Organization* (Continuum International Publishing, 2009)

Fortescue, Michael, 'Similitude: A Conceptual Category' (2010) 42(2) *Acta Linguistica Hafniensia* 117

Fox, Mark C, K Anders Ericsson and Ryan Best, 'Do Procedures for Verbal Reporting of Thinking Have to Be Reactive? A Meta-Analysis and Recommendations for Best Reporting Methods,' (2011) 137(2) *Psychological Bulletin* 316

Freeman, J, 'If You Can't Measure It – It doesn't Exist' (2007) 18 *High Ability Studies* 65

Galton, F, *Hereditary Genius* (Macmillan, 1869)

Gantt, Larry O, 'Deconstructing Thinking Like a Lawyer: Analyzing the Cognitive Components of the Analytical Mind' (2006-2007) 29 *Campbell Law Review* 413

Gardner, H, 'Expert Performance: Its Structure and Acquisition' (1995) 50 *American Psychologist* 802

Garicano, Luis and Thomas N Hubbard, 'Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyer's Human Capital' (2007) 50(1) *Journal of Law and Economics* 1

Garicano, Luis and Thomas N Hubbard, 'Specialization, Firms, and Markets: The Division of Labor within and between Law Firms' (2008) 25(2) *The Journal of Law, Economics, & Organization* 339

Garner, Bryan A, 'Although IRAC Works for Exams, Avoid It in Practice' (2004-2005) 33 *Student Lawyer* 10

Gigerenzer, G and H Brighton, 'Homo Heuristicus: Why Biased Minds Make Better Inferences' (2009) 1 *Topics in Cognitive Science* 107

Gilbert, D T, 'Thinking Lightly About Others: Automatic Components of the Social Inference Process' in J Uleman and J A Bargh (eds), *Unintended Thought* (Prentice-Hall, 1989) 189

Gilbert, D T, 'Inferential Correction' in T Gilovich, D Griffin and D Kahneman (eds) *Heuristics and Biases* (Cambridge University Press, 2002) 167

Gilhooly, K J, P McGeorge, J Hunter, J M Rawles, I K Kirby, C Green and V Wynn, 'Biomedical Knowledge in Diagnostic Thinking: The Case of Electrocardiagram (ECG) Interpretation' (1997) 9 *European Journal of Cognitive Psychology* 199

Gladwell, Malcolm, *Outliers: The Story of Success* (Little, Brown and Company, 2008)

Glaser, Robert, 'Education and Thinking: The Role of Knowledge' (1984) 39 *American Psychologist* 93

Glockner, Andreas and Irena D Ebert, 'Legal Intuition and Expertise,' in Marta Sinclair (ed), *Handbook of Intuition Research* (Edward Elgar Publishing Limited, 2011) 157

Gobet, Fernand and Neil Charness,'Expertise in Chess,' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 523

Goor, A and R E Sommerfeld, 'A Comparison of Problem-Solving Processes of Creative and Non-Creative Students' (1975) 67 *Journal of Educational Psychology* 495

Gordon, Leslie A, *Legal Wranglers: Law Firms Add Coaches to their Staffs*, ABA Journal, May 2014

Grant, Kenneth A, 'Tacit Knowledge Revisited – We Can Still Learn from Polanyi' (2007) 5(2) *The Electronic Journal of Knowledge Management* 173

Greaves, C E, H Zacher, B McKenna and D Rooney, 'Wisdom and Narcissism as Predictors of Transformational Leadership' (2014) 35(4) *Leadership & Organization Development Journal* 335

Groen, C J and Vimla L Patel, 'Medical Problem-Solving: Some Questionable Assumptions' 1985) 19(2) *Medical Education* 95

Hambrick, David Z, Frederick L Oswald, Erik M Altmann, Elizabeth J Meinz, Fernand Gobet and Guillermo Campitelli, 'Deliberate Practice: Is That All It Takes To Become An Expert?' (2013) 45 *Intelligence* 34

Hammond, K R, *Judgment Under Stress* (Oxford University Press, 2000)

Hayes, J R, *The Complete Problem Solver* (Franklin Institute Press, 1981)

Henderson, Bethany Rubin, 'Asking the Lost Question: What is the Purpose of Law School?' (2003) 53(1) *Journal of Legal Education* 48

Henderson, William D, *Supercharging Lawyer Development Through Feedback*, NALP Bulletin, June 2014

Hinds, P J, 'The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Prediction of Novice Performance' (1999) 5 *Journal of Experimental Psychology: Applied* 205

Hoffman, R R, 'The Problem of Extracting the Knowledge of Experts from the Perspective of Experimental Psychology' (1987) 8 *AI Magazine* 53

Hoffman, Robert R, 'How Can Expertise be Defined? Implications of Research from Cognitive Psychology' in Robin Williams, Wendy Faulkner and James Fleck (eds), *Exploring Expertise: Issues and Perspectives* (Macmillan Press Limited, 1998) 81

Hoffman R R, J W Coffey and K M Ford, *A Case Study in the Research Paradigm of Human-Centered Computing: Local Expertise in Weather Forecasting* (National Technology Alliance, 2000)

Hoffman, R R, G Trafton and P Roebber, *Minding the Weather: How Expert Forecasters Think* (MIT Press, 2005)

Hoffman, Robert R, Paul Ward, Paul J Feltovich, Lia DiBello, Stephen M Fiore and Dee H Andrew, *Accelerated Expertise: Training for High Proficiency in a Complex World* (Psychology Press, 2014)

Hogarth, R M, *Educating Intuition* (University of Chicago Press, 2001)

International Competition Network, *ICN Investigative Techniques Handbook for Merger Review* (International Competition Network, June 2005)

International Competition Network, *ICN Recommended Practices for Merger Analysis* (International Competition Network, 2008)

Ippolito, Richard A, 'The Sorting Function: Evidence from Law School' (2001) 51 *Journal of Legal Education* 533

Jacoby, L L, 'A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory (1991) 30 *Journal of Memory and Language* 513

Jacoby, L L, 'Dissociating Automatic and Consciously Controlled Effects of Study/Test Compatibility' (1996) 35 *Journal of Memory and Language* 32

James, Nickolas, Clair Hughes and Clare Cappa, 'Conceptualising, Developing and Assessing Critical Thinking in Law' (2010) 15(3) *Teaching in Higher Education* 285

James, Nick, *Good Practice Guide (Bachelor of Laws): Thinking Skills (Threshold Learning Outcome 3)* (Australian Learning and Teaching Council, 2011)

James, Nickolas John, 'Professional Mentoring Programs for Law Students' (2011) 30 *University of Tasmania Law Review* 90

James, Nick, 'Logical, Critical and Creative: Teaching 'Thinking Skills' to Law Students' (2012) 12(1) *QUT Law & Justice* 66

Jenkins, J J, 'Educating for Applications' in R R Hoffman and D S Palermo (eds), *Cognition and the Symbolic Processes: Applied and Ecological Perspectives* (Erlbaum, 1983) 335

Johnson, E S 'An Information Processing Model of One Kind of Problem Solving' (1964) 4 *Psychological Monographs* 78

Johnson, P E, 'The Expert Mind: A New Challenge for the Information Scientist,' in TMA Bemelmans (ed), *Beyond Productivity: Information Systems Development for Organization Effectiveness* (North Holland Publishing, 1984) 367

Johnson, E J, 'Expertise and Decision Under Uncertainty: Performance and Process,' in MTH Chi, R Glaser and MJ Farr (eds) *The Nature of Expertise* (Erlbaum, 1988) 209

Justice Gault, 'A Judicial Perspective on Competition Litigation,' in R Adhar (ed), *Competition Law and Policy in New Zealand* (Law Book Co., 1991) 96

Kahn, Larry, *Think Like a Lawyer: Negotiating a New Car Purchase* (Amazon Digital Services, 2012)

Kahneman, D and A Tversky, 'On the Study of Statistical Intuition' in D Kahneman, P Slovic and A Tversky (eds), *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press, 1982) 493

Kahneman, D and S Frederick, 'Representativeness Revisited: Attribute Substitution in Intuitive Judgment' in T Gilovich, D Griffin and D Kahneman (eds), *Heuristics and Biases* (Cambridge University Press, 2002) 49

Kahneman, Daniel, 'A Perspective on Judgment and Choice: Mapping Bounded Rationality' (2003) 58(9) *American Psychologist* 697

Kahneman, Daniel, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011)

Kallaugher, J, 'Review of Elhauge & Geradin's Global Competition Law and Economics' (2007) 3(2) *Competition Policy International* 241

Kaufman, S B, 'Investigating the Role of Domain General Mechanisms in the Acquisition of Domain Specific Expertise' (2007) 18 *High Ability Studies* 71

Kedia, Soma, 'Redirecting the Scope of First-Year Writing Courses: Towards a New Paradigm of Teaching Legal Writing' (2009-2010) 87 *University of Detroit Mercy Law Review* 147

Kermarrec, Gilles and Cyril Bossard, 'Defensive Soccer Players' Decision Making: A Naturalistic Study' (2014) 8(2) *Journal of Cognitive Engineering and Decision Making* 187

Keyzer, Patrick, *Legal Problem Solving – A Guide for Law Students* (LexisNexis, 2002)

Kift, Sally, Mark Israel and Rachael Field, *Learning and Teaching Academic Standards Project: Bachelor of Laws Learning and Teaching Academic Standards Statement, December 2010* (Australian Learning and Teaching Council, 2010)

King, L and N Tuckwell, *Stimulated Recall Memory* (The Western Australian Institute for Educational Research, 1983)

Klein, G A, 'A Recognition Primed Decision (RPD) Model of Rapid Decision Making,' in G A Klein, J Orasanu, R Calderwood and C E Zsambok (eds), *Decision-Making in Action: Models and Methods* (Ablex, 1993) 138

Klein, G, *Sources of Power: How People Make Decisions* (MIT Press, 1998)

Klein, D E, H A Klein and G Klein, 'Macrognition: Linking Cognitive Psychology and Cognitive Ergonomics,' in *Proceedings of the 5th International Conference on Human Interactions with Complex Systems* (University of Illinois, 2000) 173

Klein, G, *Intuition at Work: Why Developing Your Gut Instincts Will Make You Better at What You Do* (Doubleday, 2003)

Klein, Gary, Karol G Ross, Brian M Moon, Devorah E Klein, Robert R Hoffman and Erik Hoffnagel, 'Macrognition' (2003) 3 *Human-Centered Computing* 81

Klein, Gary, 'Naturalistic Decision Making' (2008) 50(3) *Human Factors* 456

Klein, Gary, *Seeing What Others Don't: The Remarkable Ways We Gain Insights* (Public Affairs, 2013)

Klinger, E, 'Utterances to Evaluate Steps and Control Attention Distinguish Operant from Respondent Thought While Thinking Out Loud' (1974) 4 *Bulletin of the Psychonomic Society* 44

Krieger, Stefan H, 'Domain Knowledge and the Teaching of Creative Legal Problem Solving' (2004-2005) 11 *Clinical Law Review* 149

Krieger, Stefan H, 'The Development of Legal Reasoning Skills in Law Students: An Empirical Study' (2006) 56(3) *Journal of Legal Education* 332

Lakatos, I, *Proofs and Reputations* (Cambridge University Press, 1976)

Larkin, J H, J McDermott, D P Simon and H A Simon, 'Models of Competence in Solving Physics Problems' (1980) 4 *Cognitive Science* 317

Lawrence, Jeanette A, 'Expertise on the Bench: Modeling Magistrate's Judicial Decision-Making' in Michelene T H Chi, Robert Glaser and Marshall J Farr (eds), *The Nature of Expertise* (Lawrence Erlbaum Associates, 1988)

Leahey, Erin and Laura A Hunter, 'Lawyer's Lines of Work: Specialization's Role in the Income Determination Process' (2012) 90(4) *Social Forces* 1101

Le Brun, Marlene and Richard Johnstone, *The Quiet (R)Evolution: Improving Student Learning in Law* (Law Book Co, 1995)

Lemaire, P and R S Siegler, 'Four Aspects of Strategic Change: Contributions to Children's Learning of Multiplication' (1995) 124 *Journal of Experimental Psychology* 83

Lesgold, A, H Rubinson, P Feltovich, R Glaser, D Klopfer and Y Wang, 'Expertise in a Complex Skill: Diagnosing X-ray Pictures,' in Michelene T H Chi, Robert Glaser and Marshall J Farr (eds), *The Nature of Expertise* (Erlbaum, 1988) 311

Levelt, Willem J M, Ardi Roelofs and Antje s Meyer, 'A Theory of Lexical Access in Speech Production' (1999) 22 *Behavioral and Brain Sciences* 1

Levitin, Daniel J, *This is Your Brain on Music: The Science of a Human Obsession* (Plume/Penguin, 2007)

Lipshitz, Raanan, Gary Klein, Judith Orasanu and Eduardo Salas, 'Tacking Stock of Naturalistic Decision Making' (2001) 14(5) *Journal of Behavioural Decision Making* 331

Logie, R H, 'Visuo-Spatial Processing in Working Memory' (1986) 23A *Quarterly Journal of Experimental Psychology* 229

Lundeberg, Mary A, *Studying Understanding in Legal Case Analysis (Reading, Self-Control Training, Metacognition)* (PhD Thesis, University of Minnesota, 1985)

Lundeberg, Mary A, 'Metacognitive Aspects of Reading Comprehension: Studying Understanding in Legal Case Analysis' (1987) 22(4) *Reading Research Quarterly* 407

Lundgren-Laine, Helja and Sanna Salantra, 'Think-Aloud Technique and Protocol Analysis in Clinical Decision-Making Research' (2010) 20(4) *Qualitative Health Research* 565

Maister, David, *Managing The Professional Service Firm* (Free Press, 1997)

Makri, Stephann, Ann Blandford and Anna L Cox, 'This is What I am Doing and Why: Methodological Reflections on a Naturalistic Think-Aloud Study of Interactive Information Behaviour' (2011) 47 *Information Processing and Management* 336

Malterud, K, 'Qualitative Research: Standards, Challenges, and Guidelines' (2001) *Lancet* 483

Marchant, Garry, John Robinson, Urton Anderson and Michael Schadewald, 'Analogical Transfer and Expertise in Legal Reasoning' (1991) 48 *Organizational Behaviour and Human Decision Processes* 272

McKinney, Ruth Ann, *Reading Like A Lawyer: Time-Saving Strategies for Reading Like an Expert* (Carolina Academic Press, 2nd ed, 2012)

McNeilage, Amy, 'Universities Investigate MyMaster Cheating,' *Sydney Morning Herald*, 18 November 2014

Metzler, Jeffery, 'The Importance of IRAC and Legal Writing' (2002-2003) 60 *University of Detroit Mercy Law Review* 501

Meyer, Antje, Linda Wheeldon and Mark Smith, *Encyclopedia of Cognitive Science* (John Wiley & Sons, 2005)

Middleton, H E, *The Role of Visual Mental Imagery in Solving Complex Problems in Design* (PhD Thesis, Griffith University, 1998)

Mitchell, John B, 'Current Theories on Expert and Novice Thinking: A Full Faculty Considers the Implications for Legal Education' (1989) 39 *Journal of Legal Education* 275

Montgomery, H and C M Allwood, 'On the Subjective Representation of Statistical Problems' (1978) 22 *Scandinavian Journal of Educational Research* 107

Morris, Clarence, *How Lawyers Think* (Harvard University Press, 1937)

Moxely, Jerad H, K Anders Ericsson, Neil Charness and Ralf T Krampe, 'The Role of Intuition and Deliberative Thinking in Experts' Superior Tactical Decision-Making' (2012) 124 *Cognition* 72

Myers, D G, *Intuition: Its Powers and Perils* (Yale University Press, 2002)

Newell, Allen and Herbert A Simon, *Human Problem Solving* (Prentice Hall, 1972)

New Zealand Commerce Commission, *Mergers and Acquisitions Guidelines* (Commerce Commission, July 2013)

Nievelstein, Fleurie, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Expertise-Related Differences in Conceptual and Ontological Knowledge in the Legal Domain' (2008) 20(6) *European Journal of Cognitive Psychology* 1043

Nievelstein, Fleurie, Tamara van Gog, Henny P A Boshuizen and Frans J Prins, 'Effects of Conceptual Knowledge and Availability of Information Sources on Law Student's Legal Reasoning' (2010) 38 *Instructional Science* 23

Nievelstein, Fleurie, Tamara van Gog, Gijs van Dijck and Henny P A Boshuizen, 'Instructional Support for Novice Law Students: Reducing Search Processes and Explaining Concepts in Cases' (2011) 25 *Applied Cognitive Psychology* 408

Nodine, C F, H L Kundel, S C Lauver and L C Toto, 'Nature of Expertise in Searching Mammograms for Breast Masses' (1996) 3 *Academic Radiology* 1000

Nodine, Calvin F, Harold L Kundel, Claudia Mello-Thoms, Susan P Weinstein, Susan G Orel, Daniel C Sullivan and Emily F Conant, 'How Experience and Training Influence Mammography Expertise' (1999) 6(10) *Academic Radiology* 575

Norman, Geoffrey R, Donald Rosenthal, Lee R Brooks, Scott W Allen and Linda J Muzzin, 'The Development of Expertise in Dermatology' (1989) 125 *Archives of Dermatology* 1063

Oates, Laurel Currie, 'Beating the Odds: Reading Strategies of Law Students Admitted Through Alternative Admissions Programs' (1997-1998) 83 *Iowa Law Review* 139

Oates, Laurel Currie, 'Leveling the Playing Field: Helping Students Succeed by Helping Them Learn to Read as Expert Lawyers' (2006) 80 *St John's Law Review* 227

Ohlsson, S, *The Cognitive Seminar: Report No 6 – Competence and Strategy in Reasoning with Common Spatial Concepts* (University of Stockholm, 1980)

Okamoto, Karl S, 'Teaching Transactional Lawyering' (2009) 1 *Drexel Law Review* 69

Oskamp, S, 'Overconfidence in Case-Study Judgments' (1965) 29 *Journal of Consulting Psychology* 261

Palasota, Anthony, 'Expertise and the Law: Some Recent Findings from the Cognitive Sciences About Complex Human Information Processing' (1990-1991) 16 *Thurgood Marshall Law Review* 599

Pasupathi, M, U M Staudinger and P B Baltes, 'Seeds of Wisdom: Adolescent's Knowledge and Judgment About Difficult Life Problems' (2001) 37 *Developmental Psychology* 351

Patel, Vimla, G J Groen and C H Frederiksen, 'Differences Between Medical Students and Doctors in Memory for Clinical Cases' (1986) 20(1) *Medical Education* 3

Patel, V L, G J Groen and J F Arocha, 'Medical Expertise as a Function of Task Difficulty' (1990) 18(4) *Memory & Cognition* 394

Patel, Vimla L and Guy J Groen, 'The General and Specific Nature of Medical Expertise: A Critical Look' in K Anders Ericsson and Jacqui Smith (eds), *Toward a General Theory of Expertise: Prospects and Limits* (Cambridge University Press, 1991) 93

Patel, V L and G J Groen, 'Developmental Accounts of the Transition from Medical Student to Doctor: Some Problems and Suggestions' (1991) 25(6) *Medical Education* 527

Patel, V L and D R Kaufman, 'Clinical Reasoning and Biomedical Knowledge: Implications for Teaching,' in J Higgs and M Jones (eds) *Clinical Reasoning in the Health Professions* (Butterworth-Heinemann Ltd, 1995) 117

Pemberton, John De J, 'National Law Student Conference: The Conference Report,' (1948-49) 1 *Journal of Legal Education* 73

Perkins, David N, 'Reasoning as Imagination' (1985) 16 *Interchange* 14

Pfaff, Mark S, Gary L Klein, Jill L Drury, Sung Pil Moon, Yikun Liu and Steven O Entezari, 'Supporting Complex Decision Making Through Option Awareness' (2013) 7(2) *Journal of Cognitive Engineering and Decision Making* 155

Pink, Daniel H, *Drive: The Surprising Truth About What Motivates Us* (Riverhead Books, 2011)

Polanyi, Michael, *The Tacit Dimension* (Doubleday, 1966)

Polanyi, Michael, *The Tacit Dimension* (Doubleday, 1973)

Polanyi, Michael, *Personal Knowledge: Towards a Post-Critical Philosophy* (University of Chicago Press, 1974)

Popper, K, *The Logic of Scientific Discovery* (Hutchison, 1959)

Posner, Richard, *How Judges Think* (Harvard University Press, 2008)

Raab, M and J G Johnson, 'Expertise-Based Differences in Search and Option-Generation Strategies' (2007) 13 *Journal of Experimental Psychology: Applied* 158

Rapoport, Nancy B, 'Is "Thinking Like a Lawyer" Really What We Want to Teach?' (2002) 1 *Journal of the Association of Legal Writing Directors* 91

Rodriguez, Sheila, 'Using Feedback Theory to Help Novice Legal Writers Develop Expertise' (2008-2009) 86 *University of Detroit Mercy Law Review* 207

Romantz, David S and Kathleen Elliot Vinson, *Legal Analysis: The Fundamental Skill* (Carolina Academic Press, 2nd ed, 2009)

Ross, Karol G, Jennifer L Shafer and Gary Klein, 'Professional Judgments and "Naturalistic Decision Making,"' in K Anders Ericsson, Neil Charness, Paul J Feltovich and Robert R Hoffman (eds), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge University Press, 2006) 403

Round, D, 'Market Definition – A Matter of Fact, Commercial Commonsense or Economic Principles?' in A Bollard (ed), *The Economics of the Commerce Act* (New Zealand Institute of Economic Research, 1989) 20

Saariluomo, P, 'Apperception and Restructuring in Chess Players' Problem Solving,' in K Gilhooly, M T G Keane, R H Logie and G Erdos (eds), *Lines of Thinking: Reflections of the Psychology of Thought* (John Wiley & Sons, 1990)

Sanson, Michelle, 'Thinking Like a Lawyer' (2006) *International Bar Association Conference Newsletter*

Sanson, Michelle, Thalia Anthony and David Worswick, *Connecting with Law* (Oxford University Press, 2nd ed, 2010)

Sargent, S S, 'Thinking Processes at Various Levels of Difficulty' (1940) *Archives of Psychology* 249

Scandura, J M, 'Problem Solving in Schools and Beyond: Transitions from the Naïve to the Neophyte to the Master' (1981) 16(3) *Educational Psychologist* 139

Schauer, Frederick, *Thinking Like a Lawyer: A New Introduction to Legal Reasoning* (Harvard University Press, 2009, reprint 2012)

Schmidt, H G, G R Norman and H P A Boshuizen, 'A Cognitive Perspective on Medical Expertise: Theory and Implications' (1990) 65(10) *Academic Medicine* 611

Schmidt, H G and P A Boshuizen, 'On Acquiring Expertise in Medicine' (1993) 5 *Educational Psychology Review* 205

Schneider, W, 'Training High Performance Skills: Fallacies and Guidelines' (1985) 27(3) *Human Factors* 285

Schraagen, Jan Maarten, 'How Experts Solve a Novel Problem in Experimental Design' (1993) 17(2) *Cognitive Science* 285

Schwalbe, Ulrich and Daniel Zimmer, *Law and Economics in European Merger Control* (Oxford University Press, 2009)

Schweickert, R, A M Burton, N K Taylor, E N Corlett, N R Shadbolt and A P Hedgecock, 'Comparing Knowledge Elicitation Techniques: A Case Study' (1987) 1 *Artificial Intelligence Review* 245

Senger, Charles J, *Learning Legal Reasoning in Law School: The Differences Between First and Third Year Students* (PhD Thesis, Michigan State University, 1989)

Senger, Charles J, 'Thinking Aloud Protocols: A Diagnostic Tool for Teaching Legal Problem Solving' (1993) 10 *Thomas M Cooley Law Review* 367

Shadbolt, N R and A M Burton, 'Kowledge Elicitation Elicitation Techniques: Some Experimental Results,' in KL McGraw and CR Westphal, *Readings in Knowledge Acquisition* (Ellis Horwood,1990)

Shenk, David, *The Genius in All of Us: Why Everything You've Been Told about Genetics, Talent and IQ is Wrong* (Image Books, 2010)

Sherr, Avrom, 'The Value of Experience in Legal Competence' (2000) 7(2) *International Journal of the Legal Profession* 95

Simon, D P and H A Simon, 'Individual Differences in Solving Physics Problems,' in R Siegler (ed) *Children's Thinking: What Develops?* (Erlbaum, 1978) 325

Simon, H A and W G Chase, 'Skill in Chess' (1973) 61 *American Scientist* 394

Smith, C O, *The Structure of Intellect Processes Analyses Systems: A Technique for the Investigation and Qualification of Problem Solving Processes* (PhD Thesis, University of Houston, 1971)

Smith, J and P B Baltes, 'Wisdom-Related Knowledge: Age/Cohort Differences in Response to Life-Planning Problems,' (1990) 26 *Developmental Psychology* 494

Smith, J, U M Staudinger and P B Baltes, 'Occupational Settings Facilitating Wisdom-Related Knowledge: The Sample Case of Clinical Psychologists' (1994) 62 *Journal of Consulting and Clinical Psychology* 989

Smith, R and J Walker, 'Australian Trade Practices and the Emerging Role of "Commercial Reality" Versus Substitution in Market Definition' (1997) 5 *Competition & Consumer Law Journal* 1

Spence, Michael, 'Job Market Signaling' (1973) 87(3) *The Quaterly Journal of Economics* 355

Stanovich, Keith E and Richard F West, 'Individual Differences in Reasoning: Implications for the Rationality Debate?' (2000) 23 *Behavioral and Brain Sciences* 645

Stanovich, K E and R F West, 'Individual Differences in Reasoning: Implications for the Rationality Debate' in T Gilovich, D Griffin and D Kahneman (eds), *Heuristics and Biases* (Cambridge University Press, 2002) 421

Staudinger, U M, J Smith and P B Baltes, 'Wisdom-Related Knowledge in a Life Review Task: Age Differences and the Role of Professional Specialization' (1992) 7 *Psychology and Aging* 271

Staudinger, Ursula M, Jacqui Smith and Paul B Baltes, *Manual for the Assessment of Wisdom-Related Knowledge* (Max Planck Institute for Human Development and Education, 1994)

Sternberg, Robert J, 'Words to the Wise About Wisdom: A Commentary of Ardelt's Crtique of Baltes' (2004) 47 *Human Development* 286

Sternberg, Robert J, 'What Do We Know About Tacit Knowledge? Making the Tacit Become Explicit,' in Robert J Sternberg and Joseph A Horvath (eds), *Tacit Knowledge in Professional Practice: Researcher and Practitioner Perspectives* (Lawrence Erlbaum Associates, 1999)

Sternberg, R J and P A Frensch, 'On Being an Expert: A Cost Benefit Analysis,' in R R Hoffman (ed), *The Psychology of Expertise: Cognitive Research and Empirical AI* (Springer Verlag, 1992) 191

Stratman, James F, 'Contract Disclaimers in ERISA Summary Plan Documents: A Deceptive Practice?' (1988) 10 *Industrial Relations Law Journal* 350

Stratman, James F, 'The Emergence of Legal Composition as a Field of Inquiry: Evaluating the Prospects' (1990) 60(2) *Review of Educational Research* 153

Stratman, James F, 'Adversarial and Scholarly Theories of the Appellate Court Brief: Implications for Appeal Court Judges and Their Staff Attorney's as Audiences,' in D Roen and G Kirsch (eds), *A Sense of Audience in Written Communication* (Sage, 1990) 115

Stratman, James F, 'Investigating Persuasive Processes in Legal Discourse in Real Time: Cognitive Biases and Rhetorical Strategy in Appeal Court Briefs' (1994) 17 *Discourse Processes* 1

Stratman, James F, 'When Law Students Read Cases: Exploring Relations Between Professional Legal Reasoning Roles and Problem Detection' (2002) 34(1) *Discourse Processses* 57

Stratman, James F, 'How Legal Analysts Negotiate Indeterminacy of Meaning in Common Law Rules: Toward a Synthesis of Linguistic and Cognitive Approaches to Investigation' (2004) 24 *Language & Communication* 23

Stromso, Helge I and Ivar Braten, 'Norwegian Law Students' Use of Multiple Sources While Reading Expository Texts' (2002) 37(2) *Reading Research Quarterly* 208

Susskind, Richard, *The Future of Law: Facing the Challenges of Information Technology* (Oxford University Press, 1998)

Susskind, Richard, *The End of Lawyers? Rethinking the Nature of Legal Services* (Oxford University Press, 2010)

Susskind, Richard, *Tomorrow's Lawyers: An Introduction to Your Future* (Oxford University Press, 2013)

Syed, Matthew, *Bounce: Mozart, Federer, Picasso, Beckham and the Science of Success* (Harper Perennial, 2011)

Taleb, Nassim N, *The Black Swan: The Impact of the Highly Improbable* (Penguin Books, 2007)

Taleb, Nassim N, *Antifragile: How to Live in a World We Don't Understand* (Allen Lane, 2012)

Topolinski, S and F Strack, 'The Architecture of Intuition: Fluency and Affect Determine Intuitive Judgments of Semantic and Visual Coherence and Judgments of Grammaticality in Artificial Grammar Learning' (2009) 138 *Journal of Experimental Psychology* 39

Tordesillas, R S and S Chaiken, 'Thinking Too Much of Too Little? The Effects of Introspection on the Decision-Making Process' (1999) 25 *Personality and Social Psychology Bulletin* 623

Toulmin, S, *The Philosophy of Science* (Hutchison, 1953)

Tversky, A and D Kahneman, 'Judgment Under Uncertainty: Heuristics and Biases' (1974) 185 *Science* 1124

Tversky, A and D Kahneman, 'Extensional vs Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment' (1983) 90 *Psychological Review* 293

Tversky, A and D Kahneman, 'Loss Aversion in Riskless Choice: A Reference-Dependent Model' (1991) 106 *Quarterly Journal of Economics* 1039

Van der Maas, H L J and E-J Wagenmakers, 'A Psychometric Analysis of Chess Expertise' (2005) 118 *American Journal of Psychology* 29

Vandevelde, Kenneth J, *Thinking Like a Lawyer: An Introduction to Legal Reasoning* (Westview Press, 1996)

Van Someren, M W, Y F Barnard and J A C Sandberg, *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes* (Academic Press, 1994)

Voss, James F, Telly R Greene, Timothy R Post and Barbara C Penner, 'Problem-Solving Skill in the Social Sciences,' in Gordon H Bower (ed) *The Psychology of Learning and Motivation* (New York, 1983) 165

Voss, J F, G Vesonder and H Spilich, 'Test Generation and Recall by High-Knowledge and Low-Knowledge Individuals' (1980) 19 *Journal of Verbal Learning and Verbal Behavior* 651

Wade, John, 'Meet MIRAT: Legal Reasoning Fragmented into Learnable Chunks' (1990-1991) 2 *Legal Education Review* 283

Watson, J B, 'Is Thinking Merely the Action of Language Mechanisms?' (1920) 11 *British Journal of Psychology* 87

Watson, J B, *Behaviorism* (The University of Chicago Press, 1930)

Weinstein, Ian, 'Lawyering in the State of Nature: Instinct and Automaticity in Legal Problem Solving' (1998-1999) 23 *Vermont Law Review* 1

Whish, Richard and David Bailey, *Competition Law* (Oxford University Press, 7th Edition 2012)

Wilson, T D and J W Schooler, 'Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions (1991) 60 *Journal of Personality and Social Psychology* 181

Wilson, T D, S D Hodges and S J LaFleur, 'Effects of Introspecting About Reasons: Inferring Attitudes from Accessible Thoughts' (1995) 69 *Journal of Personality and Social Psychology* 16

Witteman, Cilia L M and John H L van den Bercken, 'Intermediate Effects in Psychodiagnostic Classification' (2007) 23(1) *European Journal of Psychological Assessment* 56

Wojcik, Mark E, 'Add an E to Your IRAC' (2006-2007) 35 *Student Lawyer* 26

Yang, Shu Ching, 'Reconceptualizing Think-Aloud Methodology: Refining the Encoding and Categorizing Techniques via Contextualized Perspectives' (2003) 19 *Computers in Human Behavior* 95

B *Cases*

Bertelsman SE & Co. KGaA and Pearson plc (File Number 14007, 20 December 2012, Cleared 19 March 2013) Decision Number [2013] NZCC 6

Bligh Finance Limited and Hire Equipment Group Limited (Hirepool/Hirequip) (File Number 13913, 16 October 2012, Cleared 21 February 2013) Decision Number [2013] NZCC 2

Brambles Industries Limited – completed acquisition of CEVA Limited's Pallecon business (Australian Competition and Consumer Commission: Informal Review 50545, Date Commenced 19 March 2013, Date Completed 24 May 2013)

Perry Metal Protection Limited/ CSP Coating Systems (24 December 2012, withdrawn 24 May 2013), New Zealand Commerce Commission – Statement of Preliminary Issues, 25 January 2013.

Re Arnotts Limited v Trade Practices Commission [1990] FCA 473

Re Queensland Co-Operative Milling Association Ltd (1976) ATPR 40-012

Rocla Pty Ltd – proposed acquisition of Beresford Concrete Products Pty Ltd (Australian Competition and Consumer Commission: Informal Review 51157, Date Commenced 23 April 2013, Date Completed 19 June 2013)

Ruralco Holdings Limited – proposed acquisition of Elders Rural Services Limited (Australian Competition and Consumer Commission: Informal Review 50798, Date Commenced 19 March 2013, Date Completed 30 May 2013)

Toll Holdings Limited – proposed acquisition of Linfox Trans-Bass business (Australian Competition and Consumer Commission: Informal Review 50982, Date Commenced 22 March 2013, Date Completed 30 May 2013)

Trade Practices Commission v Australian Meat Holdings Pty Ltd (1988) 83 ALR 299

C *Legislation*

Commerce Act 1986

Competition and Consumer Act 2010

Legal Profession Act 2004

APPENDIX A – PRELIMINARY QUESTIONNAIRE

Q1 *For how many years have you been:*

(a) Working as a lawyer or economist (eg years post-admission or certification); and

(b) Working as a specialist competition law professional?

Q2 *(a) What kinds or types of competition law matters:*

(i) do you mainly work on at present; and

(ii) have you mainly worked on in the past?

(b) Approximately what percentage of your billable time over the last 12 months was spent on competition law matters?

Q3 *Have you ever worked in a competition authority? If so:*

(a) Which authority;

(b) During which period(s); and

(c) In which areas or divisions predominantly?

Q4 *(a) Have you ever undertaken post-graduate studies in competition law or competition economics? If so:*

(i) what was the course of study; and

(ii) when and where did you undertake it?

(b) Have you ever been an instructor or lecturer in such a course? If so, please provide brief details of where and when.

(c) Have you written any instructional or reference material in this area? If so, in what form? and

(d) How much time, on average, would you spend each week reading reference materials like decided court cases, annotated legislation, or topical articles written by other lawyers or academics?

Q5 *How many merger-related matters were you involved in during:*

(a) The last 12 months;

(b) The last five years; and

(c) Your professional career to date?

APPENDIX B – TEST-CASE SUMMARIES

A *Case A – Rocla/BCP*

This case concerned Rocla Pty Ltd's proposed acquisition of Beresford Concrete Products Pty Ltd ('BCP').[651] The Australian Competition and Consumer Commission ('ACCC') identified the relevant industry in which to consider the competitive effects of the transaction as 'construction.' Public review of the case began on 23 April 2013 and was completed 57 calendar days (or 27 review days[652]) later on 19 June 2013.

Rocla, a business owned by a New Zealand conglomerate, supplied a wide range of concrete infrastructure products used for constructing storm-water drainage systems, sewers, bridges, earth retention works, steel corrugated pipeline products and railway sleepers. Its customers included local government authorities, civil contractors and building developers. At the time of its application for merger clearance, Rocla operated 14 manufacturing and sales sites located in the Australian Capital Territory, New South Wales ('NSW'), Queensland ('QLD'), South Australia, Victoria and Western Australia.[653]

The target company, BCP, was a manufacturer of precast concrete products which it supplied to customers in NSW and QLD. These products were 'used for a range of applications, including for drainage, sewers, service access, environmental applications, roads, building products, metal products and recreational construction projects.'[654] BCP's customer base included large and small civil construction companies, as well as plumbers. It had three manufacturing sites in NSW and QLD.

The proposed acquisition was an asset sale consisting of the transfer of leases over BCP's three manufacturing premises at Charmhaven, Moorebank and Yatala, and the

[651] A copy of the ACCC's market inquiry letter in this case can be accessed at http://taps28fin.files.wordpress.com/2013/05/rocla-and-beresford.pdf

[652] The ACCC calculates Total Review Days as the total number of business days less public holidays and time during which the review was suspended (such as when ACCC staff requested further information from the merger parties).

[653] A certain familiarity with Australian and New Zealand geography was assumed. Geographical awareness was not an issue for study participants, who each displayed a sufficiently detailed knowledge of the locations of various states and cities mentioned in the test cases. Given the high-level nature of the assessment task, it was not considered necessary to provide maps or other reference material to any participant (nor were any requests of this nature made during testing). However, some readers may wish to familiarize themselves with appropriately-scaled maps of Australia and New Zealand.

[654] http://taps28fin.files.wordpress.com/2013/05/rocla-and-beresford.pdf

transfer of licence agreements, supply agreements, intellectual property, works in progress, outstanding orders, inventories and fixed assets.

The ACCC decided not to oppose the transaction on the grounds that it would be unlikely to substantially lessen competition 'in the QLD and NSW markets for the supply of box culverts … [or] drainage pits.'[655] It also dismissed any potential competition concerns over so-called conglomerate effects whereby the merged entity might have the opportunity to 'bundle or tie supply of box culverts and drainage pits (and/or other precast concrete products).'[656] These conclusions were based on the continuing competitive constraints from major competitors, low barriers to entry and expansion, customer countervailing power, and the pre-existing potential of another competitor to bundle similar products which had not led to any market foreclosure issues.

B *Case B – Toll/Linfox Trans-Bass*

This case concerned Toll Holdings Limited's proposed acquisition of the Linfox Trans-Bass business.[657] The ACCC identified the relevant market in which to consider the competitive effects of the transaction as 'freight forwarding.' Public review of the case began on 22 March 2013 and was completed 69 calendar days (or 38 review days) later on 30 May 2013.

Toll was described as 'a global provider of integrated logistics, employing around 45,000 people via a network of 1200 sites in 55 countries.' The focus of the review, however, was on two aspects of Toll's business. First, Toll's operation (via a joint venture with ANL Container Line) of a shipping service in the form of Bass Strait Shipping Services connecting the island state of Tasmania with mainland Australia. The cargo carried by this operation included 'road trailers, industrial mobiles and heavy lifts, trucks, light commercial vehicles and cars.' The second aspect was the company's subsidiary, Toll Tasmania, which provided freight forwarding services between Tasmania and the Mainland, and which used the services of Toll ANL.

[655] The ACCC's competition assessment and decision in this case can be accessed at http://transition.accc.gov.au/content/index.phtml/itemId/1115891/fromItemId/751043
[656] http://transition.accc.gov.au/content/index.phtml/itemId/1115891/fromItemId/751043
[657] A copy of the ACCC's market inquiry letter in this case can be accessed at http://taps28fin.files.wordpress.com/2013/05/toll-and-linfox.pdf

Linfox Trans-Bass was described as providing 'logistics services across the Asia Pacific region.' While not involved in providing any shipping service between Tasmania and mainland Australia, it did provide freight forwarding services between these places, operating both within Tasmania and between Tasmania and the Mainland centres of Melbourne and Sydney. Its customers were described as including businesses in the Tasmanian 'food and beverage industry, building industry and retailers.' The kinds of products it forwarded included 'glass containers, timber and steel, fresh produce and consumer goods.'

In its background description, the ACCC stated explicitly that both the acquirer and target in this transaction provided 'freight forwarding services between Tasmania and mainland Australia.' However, no revenue or industry figures were provided in the letter. Nor was there any mention of other competitors that may have been operating in relevant markets.

After assessing the likely competitive effects of the transaction in 'the market for the supply of freight forwarding services between Tasmania and the mainland Australia' and 'the market for the supply of Bass Strait shipping services,' the ACCC concluded there was no material competition concern in either market.[658] This was because of the strength of remaining competitors post-merger, which the ACCC believed would not be materially adversely affected by the transaction proceeding. In other words, the ACCC was satisfied that there would be sufficient competitive constraint on the merged entity to guard against price rises and reductions in service quality and availability.

C *Case C – Brambles/Pallecon*

This case concerned Brambles Industries Limited's completed acquisition of CEVA Limited's Pallecon business.[659] The ACCC identified the relevant industry in which to consider the competitive effects of the transaction as 'bulk storage / transport products.' Public review of the case began on 19 March 2013 and was completed 66 calendar days (or 38 review days) later on 24 May 2013.

[658] The ACCC's competition assessment and decision in this case can be accessed at http://transition.accc.gov.au/content/index.phtml/itemId/1115976/fromItemId/751043
[659] A copy of the ACCC's market inquiry letter in this case can be accessed at http://taps28fin.files.wordpress.com/2013/05/brambles-and-ceva.pdf

The acquirer and target in this case were given only a cursory description in the ACCC market inquiry letter. Brambles was described as 'a global provider of pallets, crates and containers using the CHEP and IFCO brands.' Pallecon was described as 'a leading provider of intermediate bulk container (IBC) solutions in Europe, Australia and New Zealand. In Australia, Pallecon trades as CEVA Logistics and supplies IBCs using the brands Pallecon, Maxicon, Haztainers and CEVA bins.' The area of primary overlap between these businesses as identified by the ACCC was 'the supply of hire IBCs.'

The transaction itself had already been completed inasmuch as it 'was part of an international acquisition in which Brambles and/or its related bodies corporate acquired all of the shares of Pallecon companies based in the United Kingdom, Germany, the Netherlands and Australia.'

The ACCC decided not to continue its investigation of this completed transaction on the basis that after identifying a 'container market' consisting of collapsible IBCs, rigid IBCs (both new and reconditioned), flexible IBCs, cardboard IBCs and drums (both new and reconditioned), there were many alternatives to the products supplied by the merger parties.[660] In this context, the ACCC characterized the market as differentiated with 'a large number of alternative suppliers offering a wide variety of substitutes' which would remain post-merger. It also dismissed concerns that Bramble's market power in the pallet segment of the industry might be used to leverage foreclosure in the market through bundling or tying strategies.

D *Case D – Ruralco*

This case concerned Ruralco Holdings Limited's proposed acquisition of Elders Rural Services Limited.[661] The ACCC identified the relevant industry in which to consider the competitive effects of the transaction as 'agricultural products.' Public review of the case began on 19 March 2013 and was completed 72 calendar days (or 39 review days) later on 30 May 2013.

[660] The ACCC's competition assessment and decision in this case can be accessed at http://transition.accc.gov.au/content/index.phtml/itemId/1115051/fromItemId/751043
[661] A copy of the ACCC's market inquiry letter in this case can be accessed at http://taps28fin.files.wordpress.com/2013/05/ruralco-and-elders.pdf

Ruralco, a publicly listed agribusiness company, was involved in the wholesaling, retailing and provision of 'expertise and services in rural merchandise, livestock agency, fertilizer, seed, wool, real estate, soft commodity risk management, grain marketing services, water solutions and water broking, finance and insurance.' It had 97 rural merchandise stores and was a wholesale supplier to 299 independent rural merchandise outlets. For the previous reporting period, its annual sales revenue was AUD1.136 billion with a net profit of tax of AUD13.8 million.

Elders was also a publicly listed company, but the target business was just one of its three operating divisions, the one concerned with rural services. This business involved the retailing of agri-products, the provision of livestock agency services, wool broking and rural real-estate services. Unlike Ruralco, Elders did not provide any wholesale services to independent agri-products retailers. It owned 213 rural merchandise stores and had annual sales of AUD1.3 billion with a net profit of AUD29.5 million.

The ACCC's assessment – which concluded with a decision not to oppose any aspect of the transaction – considered local retail markets, state-based livestock markets, regional wool-broking markets, the market for the supply of water broking and trading services, and state-based real estate agency markets.[662] In relation to the overlap of rural merchandise stores in local towns, the ACCC identified significant competitive constraints in the form of alternative suppliers in neighbouring towns and regions, and found that independent retailers would still have viable sources of wholesale supplies effectively constraining the merged entity from raising prices or diminishing services. It was similarly considered that there would continue to exist strong competitors in the markets for livestock services, wool broking services, water broking and trading services, and real estate agency services.

E *Cases E, F and G – Perry Metal, Penguin and Hirepool*

The fifth, sixth and seventh test cases in this study were clearance applications under consideration by the New Zealand Commerce Commission ('CC'). Unlike the Australian informal clearance process that operated in the first four test cases, the CC's assessment was part of a formal statutory process under section 66 of the *Commerce*

[662] The ACCC's competition assessment and decision in this case can be accessed at http://transition.accc.gov.au/content/index.phtml/itemId/1115891/fromItemId/751043

Act 1986. In all relevant substantive aspects, however, the assessment task involved the same kind of legal and applied economic analysis.

Case E concerned Perry Metal Protection Limited's ('PMP') proposed acquisition of CSP Coating Systems ('CSP').[663] The CC identified the relevant industry in which to consider the competitive effects of the transaction as 'galvanising.' There was no final determination in this case, however, as PMP's application for clearance, which was lodged by PMP on 24 December 2012, was withdrawn on 24 May 2013.[664]

The proposal, which was set out in the CC's Statement of Preliminary Issues dated 25 January 2013, was that 'PMP would acquire the Auckland and Christchurch galvanizing businesses carried on by CSP … [while at the same time undertaking] to divest assets in the South Island.'[665] The galvanizing process at the core of these businesses was 'hot-dip galvanizing [a process which] protects steel from corrosion by immersing the steel in molten zinc creating a protective coating layer.'[666]

The merger parties had argued that the CC should focus first on a North Island market where PMP operated four galvanizing plants in Auckland, Hamilton, Tauranga and Wellington, but where CSP – a division of Fletcher Steel – appears to have had only one plant, in Auckland. They then identified a South Island market, where the parties 'each own one of the only two galvanizing plants in the South Island,' both in Christchurch.[667] PMP's offer to divest assets in the South Island was ostensibly made to address any concerns the Commission may have had in this second market.

Case F concerned an application by Bertelsmann SE & Co ('Bertelsmann') and Pearson plc ('Pearson') for clearance of a jointly-owned company to be called Penguin Random House, which would acquire the retail book publishing businesses of Random House and Penguin (which companies were at the time independently owned by Bertelsmann

[663] A copy of the Commerce Commission's Statement of Preliminary Issues in this case can be accessed at http://taps28nztest.files.wordpress.com/2013/03/perry-metal-protection-ltd-and-csp-coating-systems-clearance-statement-of-preliminary-issues-25-january-2013.pdf
[664] http://www.comcom.govt.nz/business-competition/mergers-and-acquisitions/clearances/clearances-register/detail/781
[665] Commerce Commission, *Statement of Preliminary Issues: Perry Metal Protection Limited / CSP Coating Systems* (Wellington, 25 January 2013) 3.
[666] Ibid 1.
[667] Ibid 3.

and Pearson, respectively).[668] This was part of a global transaction, the Australian aspects of which were cleared by the ACCC on 8 March 2013.[669] The CC published its decision on 19 March 2013.[670]

In deciding to clear the transaction, the CC considered likely competitive effects in the areas of trade book publishing, the supply of trade book distribution services and the supply of books to retailers.[671] The CC concluded that the merged entity would be effectively constrained by existing competition from other publishers, that the merger would have a negligible effect in terms of increasing the market power of the parties in the trade book distribution services market, and that other suppliers and the countervailing power of purchasers would continue to be a constraint in the market for the supply of books.

Case G concerned Bligh Finance Limited's proposed acquisition of Hire Equipment Group Limited ('Hire Equip'), which provided general construction, building and DIY equipment hire services.[672] Bligh Finance already owned the Hirepool rental business which essentially provided the same kinds of services. Hirepool operated 46 branches and Hirequip operated 36 branches, all throughout New Zealand.

In deciding to clear this transaction, the CC concluded that 'the merged entity would be constrained in each sub-regional market by the presence of existing competitors, low barriers to entry and/or the countervailing power of large customers.'[673] This assessment was based on the view that the main area of overlap between the merger parties was in the hire of building construction and maintenance equipment.

---

[668] A copy of the Commerce Commission's Statement of Preliminary Issues in this case can be access at http://taps28nztest.files.wordpress.com/2013/03/random-house-penguin-clearance-statement-of-preliminary-issues-11-january-2013.pdf

[669] http://transition.accc.gov.au/content/index.phtml/itemId/1105246/fromItemId/751043

[670] Commerce Commission, *Determination: Bertelsmann SE & Co KGaA and Pearson plc* [2013] NZCC 6 (Wellington, 19 March 2013) available at http://www.comcom.govt.nz/business-competition/mergers-and-acquisitions/clearances/clearances-register/detail/780

[671] Ibid 4.

[672] A copy of the Commerce Commission's Statement of Preliminary Issues in this case can be accessed at http://taps28nztest.files.wordpress.com/2013/03/statement-of-preliminary-issues-hirepool-hirequip-29-october-2012.pdf

[673] Media Release: Commerce Commission clears Hirepool to purchase Hirequip, 21 February 2013. The Commerce Commission's formal decision (Commerce Commission, *Determination: Bligh Finance Limited and Hire Equipment Group Limited* [2013] NZCC 2 (Wellington, 21 February 2013)) can be accessed at http://www.comcom.govt.nz/business-competition/mergers-and-acquisitions/clearances/clearances-register/detail/780

APPENDIX C – TIME AND TASK ANALYSIS: EXAMPLE TRANSCRIPT

TIME	TIME AND TASK ANALYSIS PARTICIPANT S03 – CASE C: Brambles/Pallecon	Words/10s (wpm)	Reading & Clarifying	Identifying Issues	Synthesis
00:00 00:10	Okay, Brambles. Another market-inquiry letter … Pallecon and Brambles - sounds like it's going to be with pallets	16 (96)	●		
00:10 00:20	and I'm reading through … Yes, it's an acquisition by Brambles, which is a large, very large company … The [other] company I've never heard of … Pallecon	25 (150)	●		
00:20 00:30	I'm interested in find out what it is … Provides intermediate bulk container solutions. I don't actually know what that is … and I've never heard of any of those	28 (168)	●		
00:30 00:40	Brands … The overlap is identified as being "hiring," so we're in a market where a particular storage, type of storage mechanism is	22 (132)	●		
00:40 00:50	Hired … which I know is what Brambles does … They've already completed as part of an international acquisition … and so	19 (114)	●		
00:50 01:00	the Commission is looking at it after the facts to see if it has any issues … and I'm just looking at the dates … it's actually several months after the fact, which is interesting …	33 (198)	●		
01:00 01:10	Probably suggests that they've had a complaint because otherwise it wouldn't have done anything about it after so long. They would have known about it before then	27 (162)		●	
01:10 01:20	probably in some way … and [the] usual blurb which I don't need to look at … and this one's Davy [David Jones, the signatory to the letter] … and then I'm going to Attachment A	26 (156)	●		
01:20 01:30	Don't know who I am [in response to question 1 which asks parties to describe their business] … It's asking me a very general question about whether my business competes with anything that either of Brambles	23 (138)	●		
01:30 01:40	or Pallecon does for IBCs … what type of thing I produce, whether it's for sale or hire … so a very general question about who's involved in the market	28 (168)	●		
01:40 01:50	if we're a customer, it's asking what we acquire … doesn't sound like it's particularly directed to anything … it's asking a very general question about	24 (144)	●		
01:50 02:00	the nature of competition between the two companies, which have already merged … and switching and so on, which we're not necessarily	21 (126)	●		
02:00 02:10	going to know about other than for ourselves if we're a customer … "Please identify the competitors ..." [question 5] … So doesn't sound like it's been driven by a submission from the merger	29 (174)	●		
02:10 02:20	Entities … it's been driven by a complaint or some information, but it's all questions of a very general nature … Product areas, it's asking about	24 (144)		●	
02:20 02:30	what the products that the entities provide, what they're substitutable with … So it indicates the Commission doesn't know a lot about what it's talking about	25 (150)		●	
02:30 02:40	It's asking in various, various different ways … and it's asking about switching between hiring and buying	16 (96)	●		
02:40 02:50	the product … so it's asking really, in essence, whether there's a constraint from a customer to be able to buy something as opposed to being able to hire it	29 (174)	●		
02:50 03:00	which would act as a constraint on the merged entity … it's asking a very general question about barriers to entry for someone to come in and commence hiring	30 (180)	●		

TIME	TIME AND TASK ANALYSIS PARTICIPANT S03 – CASE C: Brambles/Pallecon	Words/10s (wpm)	Reading & Clarifying	Identifying Issues	Synthesis
	these things				
03:00 03:10	countervailing power [question 9], it's a stupid question to ask if you're asking it of a customer or competitor, really … it's something you can't get a lot of information	27 (162)	●		
03:10 03:20	about from the market … it's been written by a graduate … imports, I don't understand why imports would be relevant	19 (114)		●	
03:20 03:30	if you're considering hiring, because I can't imagine that imports - that there are people overseas who are hiring things like this into Australia,	23 (138)		●	
03:30 03:40	so that's probably a question related to whether or not sales are a constraint rather than imports. Doesn't seem to be a very sensible question … And other more general ... So	30 (180)		●	
03:40 03:50	how much is this transaction in danger? … Not at all … They're questions of a very general nature … They look like they're designed to assuage some complainant … They're being asked	29 (174)			●
03:50 04:00	some time after the fact about an international merger … The Commission probably doesn't have a lot to say about … It has in its usual fashion asserted section	27 (162)			●
04:00 04:10	50 jurisdiction as opposed to section 50 capital 'A' jurisdiction, which is what it really has in respect of mergers which don't involve an acquisition of shares in	28 (168)			●
04:10 04:20	an Australian company … and this sounds like it's a, an international transaction … That creates another whole level of protection for the entities, because the Commission actually can't do	28 (168)			●
04:20 04:30	that much about it if it falls under section 50 capital 'A' … so I would be advising the parties in half an hour that they really don't have anything to worry about, subject to finding out some more information	39 (234)			●